

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 07:02:14 ; Search time 7880 Seconds

(without alignments)
10918.264 Million cell updates/sec

Title: US-09-890-813-5

Perfect score: 1985
Sequence: 1 gcaccagacagacagagag.....ataaacattatcatatct 1985

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBml: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_ov: *
22: em_or: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_hcg_hum: *
31: em_hcg_in: *
32: em_hcg_other: *
33: em_hcg_mus: *
34: em_hcg_pln: *
35: em_hcg_rtd: *
36: em_hcg_mam: *
37: em_hcg_vtc: *
38: em_sy: *
39: em_hcgo_hum: *
40: em_hcgo_mus: *
41: em_hcgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1983	99.9	1985	6	AX180344 Sequence
2	1863.2	93.9	1953	6	AX180346 Sequence
3	1062.8	53.5	2037	8	AK073189 Oryza sat
4	798	40.2	1991	8	AK061941 Oryza sat
5	791.4	39.9	1343	8	AB042521 Oryza sat
6	787.6	39.7	1916	8	AK121930 Oryza sat
7	777.6	39.2	1983	8	AK102162 Oryza sat
8	742.2	37.4	1695	8	AK135862 Glycine m
9	707.6	35.6	1710	6	AX507049 Sequence
10	707.6	35.6	1710	8	BT000493 Arabidops
11	707.6	35.6	1971	8	AY057674 Arabidops
12	706	35.6	1980	8	ATAXLYS1 X98873 A.thaliana
13	695.6	35.0	1933	8	AY089207 Arabidops
14	678.4	34.2	1923	8	ATU62020 Arabidops
15	674.6	34.0	1658	6	AX180354 Sequence
16	674.6	34.0	1658	6	BT009484 Arabidops
17	654.6	33.0	2164	8	AY088366 Arabidops
18	555	28.0	565	6	AX180340 Sequence
19	448.6	22.6	513	6	AX180342 Sequence
20	371.2	18.7	847	6	AX180350 Sequence
21	244.2	12.3	282	6	AR247449 Sequence
22	225.2	11.3	270	6	AR247494 Sequence
23	204.2	10.3	455	6	AX180348 Sequence
24	158.2	7.0	160925	8	AP004332 Oryza sat
25	139.4	7.0	167881	8	AP004330 Oryza sat
26	139.4	6.5	10808	1	U67506 Methanococ
27	129.8	6.5	110000	6	AR271569_09 Continuation (10 O
28	129.8	5.8	9800	1	AE010670 Methanosa
29	114.4	5.6	11629	1	AE010310 Methanopy
30	111.8	5.5	148354	8	AP006529 Oryza sat
31	108.4	5.4	13644	1	AE013376 Methanosa
32	106.8	5.4	82612	8	AT13185 Arabidops
33	106.6	5.1	13494	1	AE001056 Archaeogl
34	101.8	5.1	133846	8	AC092559 Oryza sat
35	101	5.1	145290	1	AE015944 Clostridi
36	92.2	4.6	30205	1	AE015944 Clostridi
37	91.4	4.6	302960	1	AE016935 Bacteroid
38	90.4	4.6	4952	8	ATY16255 Arabidops
39	90.4	4.6	65465	8	AB007650 Arabidops
40	90.4	4.5	301442	1	AE016798 Arabidops
41	89.8	4.4	13866	1	AE004127 Vibrio ch
42	87.4	4.4	249150	1	AP005332 Arabidops
43	86.6	4.3	13315	1	AE007541 Clostridi
44	86.2	4.3	89154	8	ATAC010797 Arabidops
45	86	4.3	89154	8	ATAC010797 Arabidops

ALIGNMENTS

RESULT 1
AX180344 1985 bp DNA linear PAT 06-AUG-2001
LOCUS AX180344
DEFINITION Sequence 5 from Patent WO0146393.
ACCESSION AX180344
VERSION AX180344.1 GI:15132291
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
1 Falco, S.C., Farnodu, O.O. and Thorpe, C.J.
Aspartate Kinase

Wed Mar 24 14:12:53 2004

us-09-890-813-5.rge

Page 2

JOURNAL Patent: WO 0146393-A 5 28-JUN-2001;
E.I. DU PONT DE NEMOURS AND COMPANY (US
location/Qualifiers

```

/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

```

ORIGIN

Query Match	99.9%;	Score 1983;	DB 6;	Length 1985;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1985; Conservative	0;	Mismatches	0;	Gaps 0;

QY	1	GCACCAAGACAGAGCAGAGAGACTGGAAATGGCAATCCAGGTGGAGTCCGGCTGC	60
Db	1	GCACCAAGACAGAGCAGAGAGACTGGAAATGGCAATCCAGGTGGAGTCCGGCTGC	60
QY	61	CCGCGCCCTCGTTCCTCGTCGATACCTTCGGCGAGCTCTTGGACATGTTGAGAGACTGCCTG	120
Db	61	CCGCGCCCTCGTTCCTCGTCGATACCTTCGGCGAGCTCTTGGACATGTTGAGAGACTGCCTG	120
QY	121	CTTCGGTACCCGAACCGGGGCTTCGGGGTGCAGAGAGGTGTCATATGGTGGTCCGCACTC	180
Db	121	CTTCGGTACCCGAACCGGGGCTTCGGGGTGCAGAGAGGTGTCATATGGTGGTCCGCACTC	180
QY	181	CACACAGCCGTCCGGGCCAAGCAGACGGCGGGGACGGCGCTCTTGGGGCGCTCTGTCTT	240
Db	181	CACACAGCCGTCCGGGCCAAGCAGACGGCGGGGACGGCGCTCTTGGGGCGCTCTGTCTT	240
QY	241	OGGAGGGCTTCGGGATGGAGATTTGGGGGATACAGTCCAGCTGTGATGAAATTCGAGG	300
Db	241	OGGAGGGCTTCGGGATGGAGATTTGGGGGATACAGTCCAGCTGTGATGAAATTCGAGG	300
QY	301	GTCCTCGGTGTCTGTCCGCCGAGAGATGGCTTGAAGTGGCCGGCTCATCTCAACTTCCC	360
Db	301	GTCCTCGGTGTCTGTCCGCCGAGAGATGGCTTGAAGTGGCCGGCTCATCTCAACTTCCC	360
QY	361	CGAGAGACGCCCGCTGTCCTCTCTCCATGGGGAAAAACAACAACCTTCCTCT	420
Db	361	CGAGAGACGCCCGCTGTCCTCTCTCCATGGGGAAAAACAACAACCTTCCTCT	420
QY	421	TGCTGGAGAGAGCGACTGAGGTGTGGAGTATCCATGTTTCTGAAATTCGAAGATGGAA	480
Db	421	TGCTGGAGAGAGCGACTGAGGTGTGGAGTATCCATGTTTCTGAAATTCGAAGATGGAA	480
QY	481	TATGTCCTAAAGCCTCATATCAAGACGGTGGANTGACTTGGACTTCCAGNATCTGTAA	540
Db	481	TATGTCCTAAAGCCTCATATCAAGACGGTGGANTGACTTGGACTTCCAGNATCTGTAA	540
QY	541	TACAAAGCCTTATGAAGTGAAGCAATATGAAATAGTATCCCTATGATGAAAGAGCTGAC	600
Db	541	TACAAAGCCTTATGAAGTGAAGCAATATGAAATAGTATCCCTATGATGAAAGAGCTGAC	600
QY	601	GCTTAGAGACAGTACTACCTTGTTCATTTGGAGAAATGCAATGTCCACAGAGATTTTTC	660
Db	601	GCTTAGAGACAGTACTACCTTGTTCATTTGGAGAAATGCAATGTCCACAGAGATTTTTC	660
QY	661	TGCTTATTTGAACAAATTCGTGTCAAGGACAGGAGATGACGCAATTTGATTTATTTGTTT	720
Db	661	TGCTTATTTGAACAAATTCGTGTCAAGGACAGGAGATGACGCAATTTGATTTATTTGTTT	720
QY	721	CATTACAACCTGATGAATTTGGTATCGGATATCTTGAAGCAACTATCTCTGCTGTTTC	780
Db	721	CATTACAACCTGATGAATTTGGTATCGGATATCTTGAAGCAACTATCTCTGCTGTTTC	780
QY	781	GAAAGACTTCATCGGGACTGAGTACAGGATCCAGGATACCTGTGTGTTACTGCGTCTCT	840
Db	781	GAAAGACTTCATCGGGACTGAGTACAGGATCCAGGATACCTGTGTGTTACTGCGTCTCT	840
QY	841	TGGAGAGGGCTGGAATTCGTGCTGTAACTACTTTTAAAGCCGAGGTGGTATGACTTGAC	900
Db	841	TGGAGAGGGCTGGAATTCGTGCTGTAACTACTTTTAAAGCCGAGGTGGTATGACTTGAC	900

Qy	901	TGCTACAACTTGGTAAAGCTTGGGACTGAGAGAAATTCAGGTAATGGAAGTGTGA	960
Db	901	TGCTACAACTTGGTAAAGCTTGGGACTGAGAGAAATTCAGGTAATGGAAGTGTGA	960
Qy	961	TGATGACTTTCCTGTGATCCAAATATCTAACCCACATGCAAGACTGTTCATCTTAC	1020
Db	961	TGATGACTTTCCTGTGATCCAAATATCTAACCCACATGCAAGACTGTTCATCTTAC	1020
Qy	1021	ATTGAAAGGCCACAGAACTTGTTATTTTGGTGGTCAGGTTTGGATCCACAAATGAT	1080
Db	1021	ATTGAAAGGCCACAGAACTTGTTATTTTGGTGGTCAGGTTTGGATCCACAAATGAT	1080
Qy	1081	GAGACCTGGTAGAGAGGTGATATTCAGTTAGGGTTAAGATTGATCAACACCTAAAGC	1140
Db	1081	GAGACCTGGTAGAGAGGTGATATTCAGTTAGGGTTAAGATTGATCAACACCTAAAGC	1140
Qy	1141	TCGAGGACCCCTTATATCCAGACAAAGACATGGAATANGGTCGGTGTGATCTAATG	1200
Db	1141	TCGAGGACCCCTTATATCCAGACAAAGACATGGAATANGGTCGGTGTGATCTAATG	1200
Qy	1201	CATAGGCTCAAGTCAAAATGCACTATGTTGACATTTGAGCACTCGGATGCTTGCTCA	1260
Db	1201	CATAGGCTCAAGTCAAAATGCACTATGTTGACATTTGAGCACTCGGATGCTTGCTCA	1260
Qy	1261	GTATGGTTTTCGGCAAGGATTAAGATATTTGCTATATGAAATCTATGTATATCTGT	1320
Db	1261	GTATGGTTTTCGGCAAGGATTAAGATATTTGCTATATGAAATCTATGTATATCTGT	1320
Qy	1321	GATTTGTTGGTCAACATGAAATAGTGTGTTCTGTGCTACTGTGATCCATCAAAAGTCTG	1380
Db	1321	GATTTGTTGGTCAACATGAAATAGTGTGTTCTGTGCTACTGTGATCCATCAAAAGTCTG	1380
Qy	1381	GAGTAGGGAATGATATACAGCAGGCAATGAACTTGACACTGTATGTTGAAGGCTTGGA	1440
Db	1381	GAGTAGGGAATGATATACAGCAGGCAATGAACTTGACACTGTATGTTGAAGGCTTGGA	1440
Qy	1441	AATAGCAATTTGCTCTACTTTCAGCAGAGGGCGATATTTCACTTATCGAAATGTGGA	1500
Db	1441	AATAGCAATTTGCTCTACTTTCAGCAGAGGGCGATATTTCACTTATCGAAATGTGGA	1500
Qy	1501	GCAATGCTCTCGATACTAGAAAAGACGGGACGTGTGCTGAGGAAAAGTGGGGTTAATGT	1560
Db	1501	GCAATGCTCTCGATACTAGAAAAGACGGGACGTGTGCTGAGGAAAAGTGGGGTTAATGT	1560
Qy	1561	TCAGATGATCTCGCAAGAGCGTCAAAAGTTAACATGTGCTGATATGTCATGATAGCGA	1620
Db	1561	TCAGATGATCTCGCAAGAGCGTCAAAAGTTAACATGTGCTGATATGTCATGATAGCGA	1620
Qy	1621	TGCAAGGCACTGCTAGAAAGCCCTTCATCAGGCGTTCTTGAAAGACGATGCTTATCA	1680
Db	1621	TGCAAGGCACTGCTAGAAAGCCCTTCATCAGGCGTTCTTGAAAGACGATGCTTATCA	1680
Qy	1681	AGTCGAAGCGAGAAACCTACTCGTGGGCTGATCAACGTAGGCTTTCTGGGTCCAGGCGT	1740
Db	1681	AGTCGAAGCGAGAAACCTACTCGTGGGCTGATCAACGTAGGCTTTCTGGGTCCAGGCGT	1740
Qy	1741	GTTATCTGTTATGATTCCTCACTGCTCCATGAAAGGCAATGGGCAATGGATCTTATGATC	1800
Db	1741	GTTATCTGTTATGATTCCTCACTGCTCCATGAAAGGCAATGGGCAATGGATCTTATGATC	1800
Qy	1801	ATGTTTGGCTGAAACAAGTATGTCTTCAGAGTCTCAGGCAATGATGACTCAAACTGTGT	1860
Db	1801	ATGTTTGGCTGAAACAAGTATGTCTTCAGAGTCTCAGGCAATGATGACTCAAACTGTGT	1860
Qy	1861	TTCTGTTTGAATCTGTTTGCAGACCAAGTGAAGCTCGAGCACCGATGTCTCAACAAAT	1920
Db	1861	TTCTGTTTGAATCTGTTTGCAGACCAAGTGAAGCTCGAGCACCGATGTCTCAACAAAT	1920
Qy	1921	GGCAAGCCTGTGATATATTCACATGCTGTCTAATCAATATATATATATTAACATATCA	1980
Db	1921	GGCAAGCCTGTGATATATTCACATGCTGTCTAATCAATATATATATATTAACATATCA	1980

2

Db 1981 TATCT 1985

RESULT 2
AX180346
LOCUS AX180346 1953 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 7 from Patent WO0146393.
ACCESSION AX180346
VERSION AX180346.1 GI:15132292
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoidae; Andropogoneae; Zea.

REFERENCE
1 Falco, S.C., Farnodu, O.O. and Thorpe, C.J.
Aspartate Kinase
Patent: WO 0146393-A 7 28-JUN-2001;
JOURNAL E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
source location/Qualifiers
1.1953
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match	93.9%	Score 1863.2	DB 6	Length 1953
Best Local Similarity	98.4%	Pred. No. 0		
Matches 1941	Conservative 0	Mismatches 8	Indels 23	Gaps 5
QY	11	GAGCAGAGAGACTCGAAGATGCAATCCAGTGGATCGCTCGCGCCGCCGCTC	70	
DB	1	GAGCAGAGAGAGACTCGAAGATGCAATCCAGTGGATCGCTCGCGCCGCCGCTC	60	
QY	71	GTTCCGTGATACCTCCGCGGAGCTGAGCATGTTGAGAGACTGGCGTTCGGTACC	130	
DB	61	GTTCCGTGATACCTCCGCGGAGCTGAGCATGTTGAGAGACTGGCGTTCGGTACC	120	
QY	131	CGAACCGGCGCTCGCGATGCAAGAGGTTGCAATGGTGGTGGCCGACCTCCACGCGT	190	
DB	121	CGAACCGGCGCTCGCGATGCAAGAGGTTGCAATGGTGGTGGCCGACCTCCACGCGT	180	
QY	191	CGGCGCAAGCAAGCGGAGCGGAGCGGCGTCTTGGGCGCGCTGTTCTCGAGGCGCTC	250	
DB	181	CGGCGCAAGCAAGCGGAGCGGAGCGGCGTCTTGGGCGCGCTGTTCTCGAGGCGCTC	240	
QY	251	GCGATGAGAGGATGAGGAGATCACTCAAGCTGATGAAATTGCGGAGGCTCTCGGTG	310	
DB	241	GCGATGAGAGGATGAGGAGATCACTCAAGCTGATGAAATTGCGGAGGCTCTCGGTG	300	
QY	311	TGCTGGCGCGGAGATGGTGAAGTGGCGGCGCTCATCTGACGTTCCCGAGGAGCGC	370	
DB	301	TGCTGGCGCGGAGATGGTGAAGTGGCGGCGCTCATCTGACGTTCCCGAGGAGCGC	360	
QY	371	CCCGTGTGCTTCTCTCTGCGATGGGGAACAACAACCTTCTCTGCTGGAGAG	430	
DB	361	CCCGTGTGCTTCTCTCTGCGATGGGGAACAACAACCTTCTCTGCTGGAGAG	420	
QY	431	AAGCAGTAGGAGTGGAGTATCATGTTTCTGAAATCGAAGAGTGAATGTCGAA	490	
DB	421	AAGCAGTAGGAGTGGAGTATCATGTTTCTGAAATCGAAGAGTGAATGTCGAA	480	
QY	491	AGCTTACATATCAAGAGCGTGGATGAACCTTGGACCTTCCAGGATCTGTAATCAAGCTT	550	
DB	481	AGCTTACATATCAAGAGCGTGGATGAACCTTGGACCTTCCAGGATCTGTAATCAAGCTT	539	
QY	551	T-----ATGAACTGAGCAACTATTTGAAGGATGCTATGATGAAGAGCTGACGCTAG	606	
DB	540	GCTAGATGATGACCTGAGCAACTATTTGAAGGATGCTATGATGAAGAGCTGACGCTAG	599	
QY	607	GACCGAGTACTACTGTTTCAATTGGAGAAATGATGTCACACGAGATTTTTCGCTTA	666	

DB	600	GACGAGTACTACTGTTTCAATTGGAGAAATGATGTCACACGAGATTTTTCGCTTA	659	
QY	667	TTTGAACAAAATTTGGTGTCAAGGACGCGATGATGACGATTTGATATGTTGCTATTC	726	
DB	660	TTTGAACAAAATTTGGTGTCAAGGACGCGATGATGACGATTTGATATGTTGCTATTC	719	
QY	727	AACATGATATTTGGTGAATGCGATATCTTGAAGCAACCTATCTGCTTGGCGAAG	786	
DB	720	AACATGATATTTGGTGAATGCGATATCTTGAAGCAACCTATCTGCTTGGCGAAG	779	
QY	787	ACTCATGGGAGCTGATATACAGATCCAGGATACCTGTTACTGAGGTTCTTGGAA	846	
DB	780	ACTCATGGGAGCTGATATACAGATCCAGGATACCTGTTACTGAGGTTCTTGGAA	839	
QY	847	GAGCTGAAATCTGGTGTGCTGTAATCTTAAAGCCGAGGTGATGACTGACTGAC	906	
DB	840	GAGCTGAAATCTGGTGTGCTGTAATCTTAAAGCCGAGGTGATGACTGACTGAC	899	
QY	907	AACCATGGGTAAGCCTTGGAGCTGAGGAAATTCAGGTATGGAAGATGTTGATGAT	966	
DB	900	AACCATGGGTAAGCCTTGGAGCTGAGGAAATTCAGGTATGGAAGATGTTGATGAT	959	
QY	967	ACTTACTGTGATCCAAATATCTAACCAATGACGAAAGACTGTTCACTTAAATTTGA	1026	
DB	960	ACTTACTGTGATCCAAATATCTAACCAATGACGAAAGACTGTTCACTTAAATTTGA	1019	
QY	1027	AGAGGCAAGAACTTGTCTTATTTGGTGTGAGGTTTGCATCCAAATGATGAGACC	1086	
DB	1020	AGAGGCAAGAACTTGTCTTATTTGGTGTGAGGTTTGCATCCAAATGATGAGACC	1079	
QY	1087	TGCTAGAGAGGTGATATTCAGTTAGGTTAAGATTCATACACCTTAAAGCTTCAGG	1146	
DB	1080	TGCTAGAGAGGTGATATTCAGTTAGGTTAAGATTCATACACCTTAAAGCTTCAGG	1139	
QY	1147	CACCTTATACAGACAAAGAGACATGATGAAGGTTGTTGTTACTTAACTGATGAT	1206	
DB	1140	CACCTTATACAGACAAAGAGACATGATGAAGGTTGTTGTTACTTAACTGATGAT	1193	
QY	1207	GCTCAAGTCAAAATGCTACTATGTTGGAACATTTGAGACATCGGATGCTTGGTCA	1266	
DB	1194	GCTCAAGTCAAAATGCTACTATGTTGGAACATTTGAGACATCGGATGCTTGGTCA	1253	
QY	1267	TTTTCTGGCAAGGATCAGATTTTGTATATTTGAAGATCTATGATATCTGTGATG	1326	
DB	1254	TTTTCTGGCAAGGATCAGATTTTGTATATTTGAAGATCTATGATATCTGTGATG	1307	
QY	1327	TGTTGCTACAGTGAATGATGTTTCTGTGTCACTTGAATCAAAAGATCTGAGTAG	1386	
DB	1308	TGTTGCTACAGTGAATGATGTTTCTGTGTCACTTGAATCAAAAGATCTGAGTAG	1367	
QY	1387	GGAATGATATACAGCAGGCAAGTGAACATGATGTTGAGAGACTTGAGAAATAC	1446	
DB	1368	GGAATGATATACAGCAGGCAAGTGAACATGATGTTGAGAGACTTGAGAAATAC	1421	
QY	1447	AATGTTGCTACTTACAGCAGGCGATATTTCACTTATGGAATGAGAGATC	1506	
DB	1422	AATGTTGCTACTTACAGCAGGCGATATTTCACTTATGGAATGAGAGATC	1481	
QY	1507	GTCCTGATACAGAAAAGACGAGAGCTGTGAGGAAAAGGAGGATTAATGTCAGAT	1566	
DB	1482	GTCCTGATACAGAAAAGACGAGAGCTGTGAGGAAAAGGAGGATTAATGTCAGAT	1541	
QY	1567	GATCTGCAAGAGGCTTCAAGGTTACATGTGGCTGATGTGTCATGATAGGATGCAAA	1626	
DB	1542	GATCTGCAAGAGGCTTCAAGGTTACATGTGGCTGATGTGTCATGATAGGATGCAAA	1601	
QY	1627	GGAATGATAGAGGCTTCAAGGCTTCTTGAAGAGATGTCATCAAGATGCA	1686	
DB	1602	GGAATGATAGAGGCTTCAAGGCTTCTTGAAGAGATGTCATCAAGATGCA	1661	
QY	1687	AAGCAGAACTACTCGGCTGATCAAGATGAGCTTGTGCTGAGGCTGAGGCTGATTC	1746	

Db 1662 AGCGGAGAACTACTGCTG3GCTGATCACTGAGCTTGTCTGCTCCAGGCGTGTATC 1721

Qy 1747 TGTATAGAAFTCCCACTGCTGCATGAACGGGATGGGATTCATGATCATGTT 1806

Db 1722 TGTATAGAAFTCCCACTGCTGCATGAACGGGATGGGATTCATGATCATGTT 1781

Qy 1807 TGGTGAACAAAGTATGCTTCAGGTTCTGACCAATGACTGCAAACTGTGTTCTGT 1866

Db 1782 TGGTGAACAAAGTATGCTTCAGGTTCTGACCAATGACTGCAAACTGTGTTCTGT 1841

Qy 1867 TTATGAACTGTTTGACAGACACAGAGAGCTGGAGACCGATGTCACCAAGATGGCAG 1926

Db 1842 TTATGAACTGTTTGACAGACACAGAGAGCTGGAGACCGATGTCACCAAGATGGCAG 1901

Qy 1927 CCTGTATATTAATTCCTCACTGCTCTATCAATATATATATTAATTAACATTATC 1978

Db 1902 CCTGTATATTAATTCCTCACTGCTCTATCAATATATATATTAATTAACATTATC 1953

RESULT 3
AK073189
LOCUS
DEFINITION
ACCESSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AK073189 2037 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone:J033004I10, full
insert sequence.
AK073189
FLI-CDNA, CAP trapper.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriaristaceae; Oryzaceae; Oryza.

1 The Rice Full-length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Onitsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Ootomo, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurotsaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusunegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, M., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Komori, H., Miyazaki, A., Ootomo, N., Sano, H.,
Saitoh, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

2 (base 1 to 2037)
Aichi, U., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, M.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kawagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Komori, H., Kouda, M.,
Koya, S., Kurihara, C., Kurotsaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, M.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, K.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numaaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Ootomo, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, R., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,

Takawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28k full-length cDNA clones from japonica
rice.
URL: http://cdna01.dna.affrc.go.jp/cDNA/NIAS_Rice_Full-length_cDNA_Project_Team_Kikuchi,S.,Satoh,K.,Nagata,T.,Kawagashira,N.,Doi,K.,Kishimoto,N.,Yazaki,J.,Ishikawa,M.,Yamada,H.,Hotta,I.,Kojima,K.,Namiki,T.,Ohneda,E.,Yahagi,W.,Suzuki,K.,Li,C.,Ohtsuki,K.,Shishiki,T.,andYamamoto,M.
FAIS Genome Sequencing & Analysis Group; Ootomo, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Matsubara, K. and Murakami, K.
Yoshimura, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Genome Exploration Research Group in Riken
and Genome Science Laboratory in Riken
Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashidume, M., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Ishikawa-Hirozane, T., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Komori, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numaaki, R., Ootomo, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Takawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A., and Hayashizaki, Y.
Location/Qualifiers
1..2037
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cuiEier="Nipponbare"
/db_xref="taxon:39947"
/clone="J033004I10"

FEATURES
source

ORIGIN

Query Match 53.5%; Score 1062.8; DB 8; Length 2037;
Best Local Similarity 80.3%; Pred. No. 1.8e-247;
Matches 1314; Conservative 0; Mismatches 299; Indels 23; Gaps 5;

Qy 96 CTGACATGTTGAGAGACTGCGCTGCTTGGTATCCGACACCGGCGCTCGGATGCAAG 155

Db 168 CAGCTCGTATCGTGTCTCTGCGAGAGGCTCCCTCCGCGCGCGCGAGCGCGGA 227

Qy 156 GGTTCGATAGTGTGTCGCGACCTGCGACAGCTCGGCGCGCAACGACGCGCGAG 215

Db 228 GACTGAGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 267

Qy 216 ACGGCGTCTTGGGCGCGCGCTTCTCGAGAGGCTCGGAGATGAGAGATTTGGGAGTAC 275

Db 288 GCG 347

Qy 276 TGAGGTGTGTATGAAGTTGGGAGGAGTCTCGTGTGTGTCGCGCGCGAGAGTGGCTAG 335

Db 348 TGAGGTGTGTATGAAGTTGGGAGGAGTCTCGTGTGTGTCGCGCGCGAGAGTGGCTAG 407

Qy 336 TGCGCGGCTTACCTGACGTTCCCGAGAGCGCGCGCGCGCGCGCGCGCGCGCG 395

Db 408 TGCGCGGCTTACCTGACGTTCCCGAGAGCGCGCGCGCGCGCGCGCGCGCGCG 467

Qy 396 GGAACACCAACCAACCTTCTCTGCTGTGAGAGAGGAGTGGGATGGAGTTATCC 455

Db	468	GGAAGACCAACAATCTCTCTCTCTCGCTGGGGAAGAAAGGACGAGTATGGCGGCTGATCC	527
QY	456	ATGTTTCTGAAATTCGAAGAGTGGAATATGCTCAAAAGCTTCAATATCAACGCTGGATG	515
Db	528	GTTGTTTGGAAATTTGAAGAGTGGAATTTGATCCAAAGATCTACCAATTTAAACTGTGGAG	587
QY	516	AACCTGACCTTCCAAAGMATCTGTAAATACA---AGCCTTAATGAACTGAGCACTATTTG	571
Db	588	AACCTGCACTGCCAAG-ATCTGTAAATACACACTATGCTGATGATGAACCTGAGCAGCTCTG	646
QY	572	AAAGGATATGSCATATGATGAAAGAGCTGAGCGCTGAGACAGTGACTACTCTGTTTATTT	631
Db	647	AAAGGATATTCANATGATGAAAGAGCTGAACCTTGAGACACTGATCTACCTTGTTCATTT	706
QY	632	GGAGAAATGATGTCACACAGATTTTTTCTGCTTAATTTGAACAAATTCGTGTCAAGGCA	691
Db	707	GGGAGATGATGTCATCCAGATTTTTTCTGCTCTTAATTTGAACAAATTCGAGTAAGACA	766
QY	692	CGGCAATATGAGCGCATTTGATATATTTGGTTTCAATTCATGATGAAATTTGGTATATGGGAT	751
Db	767	CGGCAATATGAGCGCATTTGATATATTTGGTTTCAATTCATGATGAAATTTGGTATATGGGAT	826
QY	752	ATCTTAAGAAACAACCTATCTGCTGTTGCGAAGAGACTTCAATGGGGACTGGAATACAGAT	811
Db	827	ATCTTAAGAAACAACCTATCTGCTGTTGCGAAGAGACTTCAATGGGGACTGGAATACAGAT	886
QY	812	CCAGCAGATACCTGTTGTTTATCTGAGGTTCCCTTGGGAAAGGGCTGGAATTCGTGCTGTAAAT	871
Db	887	CCAGCAGATACCTATTTGTTTATCTGAGGTTCCCTTGGGAAAGGGCTGGAATTCGTGCTGTAAAT	946
QY	872	ACTTTAGGCCGAGGTGGTATGATGACTTGACTGACTGTCACAAACATTTGGTAAAGCCTTGGGACTG	931
Db	947	ACTTTAGGCCGAGGTGGTATGATGACTTGACTGACTGTCACAAACATTTGGTAAAGCCTTGGGACTG	1006
QY	932	AGGAAATTCAGGTAATGAGAAAGATGTATATGATATCTTTCCTTGTATCCAAATATCTAC	991
Db	1007	AGGAAATTCAGGTAATGAGAAAGATGTATATGATATCTTTCCTTGTATCCAAATATCTAC	1066
QY	992	CCACATGCAAAAGACTGTGTCATCTTAATCAATTTGAAGAGGCCACAGAACTTGCTTATTTT	1051
Db	1067	CCCAATGCAACAACACTGTGTCATCTTAATCAATTTGAAGAGGCCACAGAACTTGCTTATTTT	1126
QY	1052	GGGCTCAGGTTTGGATGCACAATCGATGAGACCTGCTGAGAGGGTATATTCAGAT	1113
Db	1127	GGGCTCAGGTTTGGATGCACAATCGATGAGACCTGCTGAGAGGGTATATTCAGAT	1188
QY	1112	AGGGTTAAGAAATTCATACAAACCTTAAGCTCCAGGCAACCTTATTAACAGACAAAGAGAC	1171
Db	1187	AGGGTTAAGAAATTCATACAAACCTTAAACCTCCAGGCAACCTTATTAACAAAGAGAGAG	1246
QY	1172	ATGATTAAGAGTCTGGTTGTAATACCTGATGAGATGAGCTCAAGTCAATATGCACTATGTTG	1233
Db	1247	ATGATTAAGAGTCTGGTTGTAATACCTGATGAGATGAGCTCAAGTCAATATGCACTATGTTG	1308
QY	1232	GACATTTGATGACCTGCGATGCTTGTGTCAGTATGTTTTCTGGCAAGGATTCAGGATTT	1291
Db	1301	GACATTTGATGACCTGCGATGCTTGTGTCAGTATGTTTTCTGGCAAGGATTCCTCTATA	1366
QY	1292	TGCTATATTTGAAGATCTATGATATCTGTGATGTTGTGTGCTACAGTGAAGTTATGTT	1351
Db	1361	T-----TGAAGATCTAGGAGTATCTGTAAATTTGTGTGCTACTAGTGAAGTTATGTT	1414
QY	1352	TCTGTGTCACTTGATTCATCAAAAGATCTGAGATGAGGAATGATACGACGCGCAATGAA	1411
Db	1415	TCTGTGTGTCTTGATTCATCAAAAGATCTGAGAGGAGGAATCTTATTCACAG-----GAA	1468
QY	1412	CTTGACCAATGATGATGAAAGCTTGAAGAAATACCAATCTTGTCTACTTCAGCAGAGG	1477
Db	1469	CTTGACCAATGATGATGAAAGCTTGAAGAAATTCGCAATGTACATCTACTTCAGCAGAGG	1522
QY	1472	GCGATATTTTCACTTATCGGAAATGTGAGCAATCTGCTGATATCTAGAAAGACGGGA	1531
Db	1529	GCAATATATTCATCTCATTTGGAAGCGTACGGGATCATCTTATATCTGGAAGAAAGGCTTT	1588

QY	1531	CGGCGCTAGAGAAAAGTGGGGTTAAAGTTCACATCATCTCGCAAGAGACGCGCAAGGTT	1541
QY	1589	CAAGTGTTAGAGAAAAGCGGGGTAAATGTCCACATCATCTCGCAAGTGCATCCAAAGTT	1648
Db	1592	AACATGTGCGCTGATATGTCATGATAGCGATGCAAGCACTCGTAGAAGCCCTTCATCAG	1651
Db	1649	AACATGTCTCTGATGTTCATGATAGCGAAGCAACAGTGCATTAAGGCCCTCCACACAG	1708
QY	1652	GGCTTCCTTGAAGACGATGTCCCTATCCACAGTCGAAACCGAACAACCTACTGTGGCGTGA	1711
Db	1709	GGCTTCCTTGAGAGACGATGTCTTGACAGAAAGTGAGAAAGAACCCCTTCACATGAGCTGA	1768
QY	1712	TCACGTAGGCTTTGC	1727
Db	1769	TCAGCTTAGTCCGATC	1784

RESULT 4	AK061941	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
	AK061941	1991 bp	mRNA	linear	PLN 24-JUL-2003			
	Oryza sativa (japonica cultivar-group)	cdna clone:001-042-D10, full insert sequence.						
	AK061941							
	AK061941.1	GI:32971959						
	FLI_CDNA; oligo-capping.							
	Oryza sativa (japonica cultivar-group)							
	Oryza sativa (japonica cultivar-group)							
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;							
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;							
	Eriocaridaceae; Oryzaceae; Oryza.							

AUTHORS

The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences, Rice Full-length cDNA Project Team, Kikuchi, S., Satch, K., Nagata, T., Kawasashira, N., Doi, K., Kishimoto, N., Yazaki, U., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, M., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Oono, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurokawa, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Naito, K., Sugiyama, A., Mizuno, K., Yokomizo, S., Nishida, U., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, U., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, U., Carminci, P., Adachi, U., Aizawa, K., Aizawa, T., Fukuda, S., Harai, A., Hashidume, W., Hayatsu, N., Imotani, K., Inli, Y., Itoh, M., Kawaga, I., Kondo, S., Kono, H., Miyazaki, A., Oosato, N., Ota, Y., Saito, R., Saeki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.

JOURNAL
TITLE
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)

MEDLINE 22752273
PUBMED 12869764

REFERENCE
2 (bases 1 to 1991)
AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,

Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Havashida, K., Havashizaki, Y., Havatsu, N., Hiratono, K., Hiraoaka, T.

Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,

Imotani, K., Ishidiki, C., Ishii, Y., Ishikawa, M.,
Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J.,
Kawamata, M.,

Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,

Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Naitoh, T., Narikawa, P., Nitkunya, J., Nishi, K., Nomura, K.

Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Ogata, Y., Ohta, V., Okawa, Y., Onari, D., Osakabe, T., Sakai, C., Sakai, T.

Osato, N., Uta, I., Uchimo, I., Ryu, K., Satoh, H., Sakai, C., Sakai, A., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,

Shinagawa, A., Shitaki, T., Shisaki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Tagami, M., Tagami-Takeda, Y.,
Suzuki, Y.,

Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,

TITLE
JOURNAL

Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
 Direct Submission
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
 This clone is one of the 28x full-length cDNA clones from japonica rice.

COMMENT

URL: http://cdna01.dna.affrc.go.jp/cDNA/
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Sato, K., Nagata, T., Kawasashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.
 PAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nishikura, U., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagata, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imochi, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ono, M., Oseko, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takeuchi-Akanita, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
 Location/Qualifiers

FEATURES

source
 1. 1991
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiVar="Nidponbare"
 /db_xref="taxon:39947"
 /clone="001-042-D10"

ORIGIN

Query Match 40.2%; Score 798; DB 8; Length 1991;
 Best Local Similarity 74.2%; Pred. No. 5e-183;
 Matches 1053; Conservative 0; Mismatches 351; Indels 15; Gaps 3;

273 ACCTCAGCGTGTGATGAAAGTTGGGGGGTCTCGTGTCTCGTGGCGCGAGATGCGT 332
 Db ACCTCAGCGTGTGATGAAAGTTGGGGGGTCTCGTGTCTCGTGGCGCGAGATGCGT 332
 315 ACCTCAGCGTGTGATGAAAGTTGGGGGGTCTCGTGTCTCGTGGCGCGAGATGCGT 374
 Db ACCTCAGCGTGTGATGAAAGTTGGGGGGTCTCGTGTCTCGTGGCGCGAGATGCGT 374
 333 AGGTGGCGGCTCATCTCGTGAAGTCCCGAGAGAGCGCGCGTGTCTCTCTGCGCA 392
 Db AGGTGGCGGCTCATCTCGTGAAGTCCCGAGAGAGCGCGCGTGTCTCTCTGCGCA 392
 375 AGGTGGCGGCTCATCTCGTGAAGTCCCGAGAGAGCGCGCGTGTCTCTCTGCGCA 454
 Db AGGTGGCGGCTCATCTCGTGAAGTCCCGAGAGAGCGCGCGTGTCTCTCTGCGCA 454
 393 TGGGGAACCAACCAACCACTTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGT 452
 Db TGGGGAACCAACCAACCACTTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGT 452
 435 TGGGGAACCAACCAACCACTTCTGATGCGTGGGGAGAAAGCGGTGTGTGTGCA 494
 Db TGGGGAACCAACCAACCACTTCTGATGCGTGGGGAGAAAGCGGTGTGTGTGCA 494
 453 TCCATGTTTCTGAATTCGAAGAGTGAATGATGCTCAAAAGCTCATATCAAGCGTGG 512
 Db TCCATGTTTCTGAATTCGAAGAGTGAATGATGCTCAAAAGCTCATATCAAGCGTGG 512
 495 CAAATGTGTGAGAGCTGTATGAGCTCACTTATTAAGAAATTCGATTTTGGACCA 554
 Db CAAATGTGTGAGAGCTGTATGAGCTCACTTATTAAGAAATTCGATTTTGGACCA 554
 513 ATGAACCTTGAGCTTCAAGNATCTGTAAATACAGCTTT---ATGAACGTGAGCACTAT 569
 Db ATGAACCTTGAGCTTCAAGNATCTGTAAATACAGCTTT---ATGAACGTGAGCACTAT 569
 555 ATCAACTTGAGATTGATGATGAGTCAATTGTTTCCGGCTTGTGCGATGAGCTTGA 614
 Db ATCAACTTGAGATTGATGATGAGTCAATTGTTTCCGGCTTGTGCGATGAGCTTGA 614
 570 TGAAGGTATGCGTATGATGAAAGAGCTGACGCGCTGAGAGAGAGAGAGAGAGAGT 629
 Db TGAAGGTATGCGTATGATGAAAGAGCTGACGCGCTGAGAGAGAGAGAGAGAGAGT 629
 615 TTAAAGGAATTGCTATGATGAAAGAGCTTACTTTAGAGCAAGAGACTATCTGTTCT 674

QY 630 TTGAGAGATGATGTCACCAAGATTTTCTGCTATTGAAACAAATTCGTGTGAC 689
 Db 675 TGGGTAATGATGATCAACAAAGATATTTCGGGCACTTTGATTAACCTGGGGTGA 734
 QY 690 CAGCGAGTATGAGCACTTTGATATTGCTTCACTGATGATGATTTGTAATGCGG 749
 Db 735 CAGCGAGTATGAGCACTTTGATATTGCTTCACTGATGATGATTTGTAATGCGG 794
 QY 750 ATATCTTGAAGCAACCTATCTGCTGTTGCGAAGAGACTTCAATGGGAGCTGAT 809
 Db 795 ATATCTTGAAGCAACCTATCTGCTGTTGCGAAGAGACTTCAATGGGAGCTGAT 854
 QY 810 ATCAAGCATATCTGTTGTTATGCTGCTTCTTGGGAAAGGCTGGAATCTGCTGTA 869
 Db 855 GTCTGCTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 914
 QY 870 CTACTTGAAGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 929
 Db 915 CCACTTGAAGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 974
 QY 930 TGAAGAAATTCAGGTATGGAAGATGATGATGATGATGATGATGATGATGATGAT 989
 Db 975 TGAAGAAATTCAGGTATGGAAGATGATGATGATGATGATGATGATGATGATGAT 1034
 QY 990 ACCGACATGCAAGAGCTGTTCCATTAACCTTGAAGAGGCGACAGAACTGCTATT 1049
 Db 1035 ATCCAAATGCAAGAGCTGTTCCATTAACCTTGAAGAGGCGACAGAACTGCTATT 1094
 QY 1050 TTGCTGCTAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1109
 Db 1095 TTGCTGCTAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1154
 QY 1110 TTGCTGCTAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1169
 Db 1155 TTGCTGCTAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1214
 QY 1170 ACATGATTAAGATCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1229
 Db 1215 ACATGATTAAGATCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1268
 QY 1230 TGGACATTTGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1289
 Db 1269 TGGACATTTGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1332
 QY 1290 TTTGCTATATTAAGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1349
 Db 1323 TTTGCTATATTAAGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1382
 QY 1350 TTTGCTATATTAAGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1409
 Db 1393 TATCTTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1442
 QY 1410 AACTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1469
 Db 1443 AACTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1502
 QY 1470 GGGGATATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1529
 Db 1503 GGGGATATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1562
 QY 1530 GACGTGCTGAGAGAAAGTGGGTTATGATGATGATGATGATGATGATGATGATGAT 1589
 Db 1563 GACGTGCTGAGAGAAAGTGGGTTATGATGATGATGATGATGATGATGATGATGAT 1622
 QY 1590 TTAACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1649
 Db 1623 TTAACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1682
 QY 1650 AGGCGTCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1688
 Db 1683 AGGCGTCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1721

LOCUS	AB042521	1343 bp	mRNA	linear	PLN 13-MAY-2000
DEFINITION	Oryza sativa mRNA for aspartate kinase, partial cds.				
ACCESSION	AB042521				
VERSION	AB042521.1	GI:7798568			
KEYWORDS					
SOURCE	Oryza sativa				
ORGANISM	Oryza sativa				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.				
AUTHORS	1 (sites)				
TITLE	Kiyota, S.				
JOURNAL	Lysine sensitive aspartate kinase from rice				
REFERENCE	Published Only in Database (2000)				
AUTHORS	2 (bases 1 to 1343)				
TITLE	Kiyota, S.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (09-MAY-2000) Seichiro Kiyota, National Institute of Agricultural Science, Laboratory of Stress Physiology; Kamonai 2-1-2, Tsuruga, Ishikawa 913-8602, Japan				
AUTHORS	(E-mail: skiyota@affrc.go.jp, Tel: +81-298-38-8382)				
FEATURES	Location/Qualifiers				
source	1..1343				
	/organism="Oryza sativa"				
	/mol_type="mRNA"				
	/db_xref="taxon:4530"				
	<1..1065				
	/codon_start=1				
	/product="aspartate kinase"				
	/protein_id="BA95630.1"				
	/db_xref="gi:7798569"				
	/translation="MSGSTRIPFAVYKIGVAKQVADPIGITTTDDPGRNDLIEAAVPAVARLHGDMVIRDPALPIVTFGLGKMKSGAVYTDREGSDLPATTTGKALGLREIOWKWDVAGVLTCDPNYIPVPAVTFPIETEEBAEALVAGVAVLHPSMRPAREGDIPIVYKNSPKAPGLITIKOREMDKVLVTSIVLNSNTVMDIYSTMLQOFGFLAKVRSIFEDGISVDCVATSEKASIVSALDPSKIMRELIIOEQLDHVEGEXKAVHLIQDQAPRITSLIGVRRSSLLIEKAFVQLRKSQVNNVMSIQGASKYNNMLIYHDEBAQCIKALHQAFFEDVLTVEVERBALHMS"				
ORIGIN					
Query Match	39.9%; Score 791.4; DB 8; Length 1343;				
Best Local Similarity	84.6%; Pred. No. 1.9e-181;				
Matches 930; Conservative	0; Mismatches 151; Indels 16; Gaps 3;				
Y	629 TTGGAGATGATGATGTCACACGAGTTTTCCTGCTTATTGGAAGAAATTTGTCGACG 688				
D	1 TTTGGGAATGATGCTACACGAGATTTTTCGCTTATTGGAAGAAATTTGTCGACG 60				
Y	689 GCACGGCACTATGACGATTTGATTTGGTTTCATTACAATGATAAATTTGGTATGCG 748				
D	61 GCACGGCACTATGACGATTTGATTTGGTTTCATTACAATGATAAATTTGGTATGCG 120				
Y	749 GATATTTGAGAGCACTATCTCTGCTGTTGGCAAGACCTTATGCGGACCTGATACG 808				
D	121 GATATTTGAGAGCACTATCTCTGCTGTTGGCAAGACCTTATGCGGACCTGATACG 180				
Y	809 GATTCAGCATACCTGTTGTTACTACGGGTTCTTGGGAAGGGCTGGAATCTGTGCTGTA 868				
D	181 GATTCAGCATACCTGTTGTTACTACGGGTTCTTGGGAAGGGCTGGAATCTGTGCTGTA 240				
Y	869 ACTACTTTAGCGCGGAGGTGTATGATCTTACTGCTCTCAACACTTTGGTAAAGCTTGGGA 928				
D	241 ACTACTTTAGCGCGGAGGTGTATGATCTTACTGCTCTCAACACTTTGGTAAAGCTTGGGA 300				
Y	929 CTGAGAGAAATTCAGGTATGAGAAAGATTGATGTTATCTTACTGATGATCAAAATTC 988				
D	301 CTGAGAGAAATTCAGGTATGAGAAAGATTGATGTTATCTTACTGATGATCAAAATTC 360				
Y	989 TACCCATGCAAAAGCTGTTCCATCTTAACTTTGAGAGGCCACAGAACTTGCTTAT 1048				

Db		36	TACCGGAATGCAACACAGCTGTGCCTAATTAAACATTGGAAGGACGACGAACCTTCTAT	420
Oy		1049	TTTGCGTCACAGTTTTCGATCCCAATGCAATGAGCCTGTAGAAGAGTGAATATGCCA	1108
Db		421	TTTTGGGCCCAAGATTTCATCCACAATCTATGAGACTGTGTGAGAAAGSTATTACCA	480
Oy		1109	GTTTAGGTTTAAGAAATTCATACMACCCCTAAAGCTCCAGGACCCTTTATTCAGAACAA	1168
Db		481	GTCAGGGTTAAAGAAATTCCTACACCCCTAAAGCTCCAGGACCCTGTACTATAACAA	540
Oy		1169	GACATGGAANAAGGTGTGTTGTAATCTACTAGATGTGTCTGAAGTCAAATGTCTATNG	1228
Db		541	GGAATGGAATA-----GGTTGTCTTAAGTACATAGTGTCTCATGCAATGTCTCTTG	594
Oy		1229	TTGGAATTTGTGAGCACTCGGATGCTTGTCAAGTATGTTTTCTGGCAAGGTTACAGT	1288
Db		595	TTGGATATATGTAAGCACTCGGATGCTTGTCCAATTTGTTTTCTGGCAAGGTTCTTCT	654
Oy		1289	ATTGTGTATATGGAAGATCTATGTAATCTGTGTGATTTGTGTCTACCAAGTAAGTTAGT	1348
Db		655	AAT------TGAAGATCTAGGGATATCTGTGATTTGTGTGCTTCTATGTAAGTTAGT	708
Oy		1349	GTTTCTGTGTCACTTGATCCATCAAAAGATGTGAGAGGAATGTAACAGCAGCAAGT	1408
Db		709	ATTCTGTGTGCTTGTATCCATCAAAAGATGTGAGAGGAATGTAATCAACAG-----	762
Oy		1409	GAACTTGAACATATGTAATGTGAAGACCTTGAAATAATGCAATTTGTGTCTACTACAGAG	1468
Db		763	GAACTTGAACATATGTAATGTGAAGGCTTGAAATAATGCAATTTGTGTCTACTACAGAG	822
Oy		1469	AGGCGCATATTTCTATCGAATGTGAGCAATGCTCTGTATCTAGATAAAAAAGCG	1528
Db		823	AGGCGCAATATATCTACTATTTGAAAAGGTGAGCGCATCTATCTTTAACTGGAAAAAGCG	882
Oy		1529	GGAGCTGTGTGCGAGGAAAAGTGGGGGTTAAATGTTCAATGATCTCGGAAGGACGTCAGAG	1588
Db		883	TTTTCAGATGTTGAGAAAAGCGGGGTAAATGTTCAATGATCTCTCGAAGGTGATCCAGAG	942
Oy		1589	GTTTAACTATGTGCTGATATGTCATGATATGCGATGCAAGACACTGTGTAGAACCCCTTCAT	1648
Db		943	GTTTAACTATGTCTGATATGTTCAATGATATGCGAGCAAGACAGTATAAAGCCCTCTCAC	10020
Oy		1649	CAGCGGTTCTTTGAAAGACGATGTCTTATCCAAAGTGGAAAGCGAAGAACTTACTGTGGGC	1708
Db		1003	CAGCGGTTCTTGGAGACGATGTCTTATCCAAAGTGGAAAGAAAGAAAGCCCTTACATGAGC	10632
Oy		1709	TGATCAAGTATGAGGCTTTC	1727
Db		1063	TGATCAAGCTTATGCTGTAC	1081

RESULT 6
AKI21930
LOCUS
DEFINITION
ORYZA SATIVA (JAPONICA CULTIVAR-GROUP) cDNA clone: J031106K22, full insert sequence.
AKI21930
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team:
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Nemiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shienkai, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Okamoto, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.

TITLE

japonica rice Science 301 (5631), 376-379 (2003)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nishikawa, J.,
Ikeda, R., Ishibiki, J., Kawamura, M., Yoshimura, A., Mura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imomani, K., Iehi, Y., Itoh, M.,
Kagawa, I., Kondo, S., Komoto, H., Miyazaki, A., Otsu, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashidume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T.,
Hori, F., Hotta, I., Iida, Y., Ikeda, R., Imamura, K.,
Imomani, K., Ishibiki, J., Iehi, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kawagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kawamura, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Komoto, H., Kouda, M.,
Koye, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Mura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y.,
Nakamura, M., Namiki, K., Narikawa, R., Nikura, J., Nishi, K.,
Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M.,
Ooka, H., Otsu, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C.,
Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K.,
Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y.,
Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M.,
Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaki-Akashi, S.,
Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K.,
Xie, Q., Yahagi, M., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J.,
Yokomizo, S., and Yoshimura, A.
Collection, mapping, and annotation of 28K full-length cDNA clones
from japonica rice

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Unpublished
3 (bases 1 to 1916)
Kikuchi, S.
Direct Submission
Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT

This clone is one of the 32K full-length cDNA clones from japonica
rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, M., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T.,
Yamamoto, M., and Nakahama, Y.
FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamura, M.,
Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M.,
Masuda, H., Mura, J., Mizuno, K., Narikawa, R., Nishikawa, J., Oka, M.,
Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M.,
Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashidume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hirooka, T., Hori, F., Iida, Y., Imamura, K., Imomani, K., Iehi, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J.,
Kishikawa-Hirozane, T., Kishimoto, N., Kouda, M., Koye, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Mura, J.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M.,

FEATURES

source

Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaki-Akashi, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A., and Hayashizaki, Y.
Location/Qualifiers
1..1916
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J033106K22"

ORIGIN

Query Match 39.7%; Score 787.6; DB 8; Length 1916;
Best Local Similarity 74.0%; Pred. No. 1,7e-180;
Matches 1058; Conservative 0; Mismatches 351; Indels 21; Gaps 4;

274 GCTCAGCGTGTGATGAGATTCCGGGGGCTCTGTGTGTGTGCGCCGAGATGCTGA 333
243 GTTCACCGTCGCCATGAAAGTTCCGGCGGTGTGTGTGTGTGCGCGAGATGCGGGA 302
334 GGTGGCGGCTCATCTCTGAAGTCCCGAGAGAGCGCCCGTGTGTGTGTGTGTGTGTGT 393
303 GGTGGCGATTCATCTCACTGCTTCCCGAGAGAGATCCCGTGTGTGTGTGTGTGTGTGT 362
394 GGGGAAACCAACCAACACCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 453
363 GGGGAAACCAACCAACACCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 422
464 CCAATTTTCTGAAATTCGAAGATGGAATATGTCGAACCAACCTATATCAAGCTGGA 513
423 GAAGCGCTGTAAATTCGAGCTCCGAGATTCGAAGAGAGAGAGAGAGAGAGAGAGAG 482
514 TGAATCTGACTTCCAAAGATCTGTATATCAACAGCTTT--ATGAATGAGCAACTAT 570
483 TGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 542
571 GAAAGATTCCTTATGATGAAAGATGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 630
543 TAAAGGTGTCTATGATGAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 602
631 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 690
603 CGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 662
691 ACGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 750
663 TCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 722
751 TATCTTGAAGCAACTATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 810
723 CATTTTGAAGCAACTATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 782
811 TCCAGCATCTCTGT 870
783 CCTCTGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
871 TACTTGAAGCAACTATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 930
843 CATTTTGAAGCAACTATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 902
931 GAGCAAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 990
903 AAGAGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962
991 CCCAGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1050
963 TCGCAACCGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1022
1051 TGGTCTCAGGTTTGTGATTCACCAATGATGATGATGATGATGATGATGATGATGATG 1110
1023 TGGTCAAGGTTTGTGATTCACCAATGATGATGATGATGATGATGATGATGATGATG 1082
1111 TAGGTTAAGATTCATCAACCTTAAGCTCCAGGACCAAGCTTATTAACGACCAAGAG 1170

Db	1083	TAAAGTTAAGAATTCATTATAACCGTGAACA	CTTGTGATCTTTATTAACATAAACAAGAGA	1142	
Oy	1171	CATGATTAANGTCTGGTGTTACTACTAGCATATGTCGTAAGTCAATGTGCATATGTT		1230	
Db	1143	TTTAGAACCAAG-----TATATTACCAAGATTTGTATTGAATCAAACATTAATATGCT		1196	
Oy	1231	GGAATTTGAGACACTCGGATGCTGTCGATATGCTTTCTGGCAAGGGTATCAGGAT		1290	
Db	1197	GGATTAAGTGAAGCACAGGATGCTGGCCAGTATAGCTTTCTGGCCAAGGTTTTCTCAAT		1256	
Oy	1291	TTGCTATATTGAAGATCTATGTATATCTTGGAATTTGTTGCTATCCAGTGAATTATGCT		1350	
Db	1257	AT-----TTGAAGATTTTAGTATCTCTGTTGATTGTGTGCTACCAAGTGAAGTCAATAT		1310	
Oy	1351	TTCTGTGTCACCTGATCCATCCAAATCATCGAAGTGAAGGAAA	CTGATPCAGAGGCAAGTA	1410	
Db	1311	ATCATTTGACACTGATGATCCATCAAACTATGCTGTCTGTAATTTGATCAGACAG-----GA		1364	
Oy	1411	ACCTGACCATGTATAGTTGAAGAGCTTGAGAAAAATGCAATTTGCTCTATCTTCACGACAG		1470	
Db	1365	ACTTGATCACTGTATGGAAGAACTTGAAAAAATTCAGATTTTCATCTACTCGACAGACAG		1424	
Oy	1471	GGGATTAATTTCACTTATTCGGAANAATGTGAGCAATTCGCTCTCTGATACTAGAAAAACGGG		1530	
Db	1425	GTCATATCATCTCACTGATAGGGAAGTACAAAGATCGATCTTAATTTCTTGAGAGAGCTTT		1484	
Oy	1531	ACGCTGCTGAGGAAAAAGTGGGGTTAATGTTCAATATGATCTCCAGAGACGCTCAAAGGT		1590	
Db	1485	CATGTTCTTACGAGAAATAGCGTTTAAGCTCAATATCTTCAACAAGGGGCAATCCAAGT		1544	
Oy	1591	TAACATGTGCTGATATGTCATGATAGCGATGCAAAAGCACTGTGMAAGCCCTTCATCA		1650	
Db	1545	GAAACATTTTCAATGTGTGCTCCATGATAGGAGGCCAACAACATGTATACAGGCCCTTCATTC		1604	
Oy	1651	GGGCTTCTTGAAGACGATGTCCTATTCACAAGTCGAAGCGGAGACCTTAC		1700	
Db	1605	AGCATTTCTTGAGAACGCTTTCTTCTTCAGAACTTGAGAGAGAGATCTTC		1654	
RESULT 7					
AKI02162					
LOCUS		1993 bp	mRNA	linear	
DEFINITION		Oryza sativa (japonica cultivar-group)	cdna clone:j033086822, full insert sequence.	AKI02162	
ACCESSION		AKI02162			
VERSION		AKI02162.1	GI:32987371		
KEYWORDS		FUL CDNA; CAP trapper.			
SOURCE		Oryza sativa (japonica cultivar-group)			
ORGANISM		Oryza sativa (japonica cultivar-group)			
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
		Eriatroidae; Oryzeae; Oryza.			
REFERENCE		1	The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yasaki,J., Ishikawa,M., Yamada,H., Ooka,H., Kotani,K., Nankai,T., Ohneda,F., Yanagi,W., Suzuki,K., Li,C., Onteki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Sasaki,Y., Tsunoda,Y., Kurokawa,T., Kodama,T., Maeda,H., Kobayashi,M., Xie,Q., Lu,M., Nishikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimizu,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN; Kawai,U., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imocant,K., Ishii,Y., Itoh,M., Kagawa,T., Kondo,S., Komori,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M., and Hayashizaki,Y.		
TITLE		Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice			

REFERENCE	AUTHORS	PUBLISHED	JOURNAL	LINE
22752273	Science 301 (5631), 376-379 (2003)			
12869764	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hizemoto, K., Hirooka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imomani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, N., Koya, S., Kuritara, C., Kurotaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, M., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Muratsu, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Nishikawa, R., Nishikawa, T., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Otsu, N., Ota, Y., Otomo, Y., Ryu, R., Satoh, H., Sakai, C., Sakai, K., Sakane, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shiba, C., Shiba, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akaiwa, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.			
	Direct Submision			
	Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of			
	Agrobiological Sciences, Department of Molecular Genetics, Head of			
	Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki			
	305-8607 Japan (E-mail:skikuchi@nias.affrc.go.jp,			
	Tel:81-29-838-7007, Fax:81-29-838-7007)			
	This clone is one of the 26k full-length cDNA clones from japonica			
	rice.			
	URL: http://cdna01.dna.affrc.go.jp/cDNA/			
	NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,			
	Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,			
	Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,			
	Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and			
	Yamamoto, M.			
	FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,			
	Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,			
	Kodama, T., Kurotaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,			
	Mizuno, K., Narikawa, R., Nishikawa, R., Nishikawa, T., Nomura, K.,			
	Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,			
	Yoshimura, A., Matsubara, K. and Murakami, K.			
	Genome Exploration Research Group in Riken Genomic Sciences Center			
	and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,			
	Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,			
	Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hizemoto, K.,			
	Hirooka, T., Hori, F., Iida, J., Imamura, K., Imomani, K., Ishii, Y.,			
	Itoh, M., Kaga, I., Kanagawa, S., Katoh, H., Kawai, J.,			
	Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,			
	Koya, S., Kuritara, C., Matsuyama, T., Miyazaki, R., Murata, M.,			
	Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N.,			
	Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,			
	Sasaki, D., Sato, K., Satoh, K., Shiba, C., Shingawa, A., Shiraki, T.,			
	Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,			
	Takaku-Akaiwa, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,			
	Yasunishi, A. and Hayashizaki, Y.			
	Location/Qualifiers			
	1. 1983			
	Organism="Oryza sativa (japonica cultivar-group)"			
	/mol_type="mRNA"			
	/cui="cui:NIPponbare"			
	/db_xref="taxon:39947"			
	/clone="U033086622"			
	FEATURES			
	source			
	1. 1983			
	ORIGIN			
	Query Match 39.2%; Score 777.6; DB 8; Length 1983;			
	Best Local Similarity 73.9%; Pred. No. 4.6e-178;			
	Matches 1048; Conservative 0; Mismatches 350; Indels 21; Gaps 4			

```

Db      320  AGCTACCCGTGATGTAAGTTTCGGCGGCTGTAGTGGCGTGGCGGACCGGATCGGG 379
Qy      333  AGGTGGCCGGCTCATCTCTGACGTTCCCGAGAGCCCGCGTGTGTCTCTGCGCA 392
Db      380  AGGTGGCCGATCTCATCTCTGACGTTCCCGAGAGCCCGCGTGTGTCTCTGCGCA 439
Qy      393  TGGGAAAACCAACCAACCTTCTCTCTGAGAGAAAGCAGTAGGTGTGAGTTA 452
Db      440  TGGGAAAACCAACCAACCTTCTCTGAGAGAAAGCAGTAGGTGTGAGTTA 499
Qy      453  TCCATGTTCTGGAATGGAAGTGAATAGTCAAAAGCCTACATCAACAGCGTGG 512
Db      500  CAATGTGTGGAGCTTATGAGCTCACCTTTATTAAGAAATTCATTTTGGACCAATG 559
Qy      513  ATGAACCTTGACCTTCCAAAGNATCTGTAAATACAGCCCTT--ATGAACCTGGAGCAACTAT 569
Db      560  ATCAACTTGGATTGGATTAGTCAATTTGTTCCGCTGTGCGATGAGCTTGACCACTTC 619
Qy      570  TGAAGGATGCGTATGATGTAAGAGCTGACGCTAGAGACAGTACTACTTGTTCAT 629
Db      620  TTAAGGAAATGCTATGATGTAAGAGCTTACTCTTAAGACAAAGACTTACTTGTTCCT 679
Qy      630  TTGAGAAATGATGTCACAGCAATTTTCTGCTTATTTGAACAAATTCGTGCAAG 689
Db      680  TCGGTGATGATGTCACAGCAATTTTCTGCTTATTTGAACAAATTCGTGCAAG 739
Qy      690  CACGGCAGTATGACGATTTATATTTGTTTCAATACATGATGAATTTGTAATCGG 749
Db      740  CACGCAAGTATGATGATTTAATAATAGCTTTATACACTGATGATTTACAAAGCTG 799
Qy      750  ATATCTTGAAGCAACCTATCTGCTGTGCGAAGACTTCAATGGGAGCTGATACAG 809
Db      800  ATATCTTGAAGCAACCTATCTGCTGTGCGAAGACTTCAATGGGAGCTGATACAG 859
Qy      810  ATCCAGCAATACCTGTTTACTGCGTTCTTGGGAAAGGCTGGAATCTGTGCTGTAA 869
Db      860  GTCCGCTATTCCTCATATGCTCACTGTTTCTCGGAAAGGATGGAATCTGTGCTGTAA 919
Qy      870  CTACTTATGCGGAGGTGTGTGATGATGCTGTACCAACCATGTAAGGCTTGGAGC 929
Db      920  CCACCTTAGGCAAGGGGTGTGTGATGATGCTGTACCAACCATGTAAGGCTTGGAGC 979
Qy      930  TGAAGAAATTCAGTATGTAAGAAATGTTGATGTGTACTTCTGTGTCAATATCT 989
Db      980  TAAAGAAATTCAGTATGTAAGAAATGTTGATGTGTACTTCTGTGTCAATATCT 1039
Qy      990  ACCGACATGCAAAAGCTGTTCATCTTAACTTTGAAGAGGCAAGAACTTGTATT 1049
Db      1040  ATCCAAATGCAAAAGCTGTTCATCTTAACTTTGAAGAGGCAAGAACTTGTATT 1099
Qy      1050  TTGTGCTCAGGTTTTCATCCAAATGATGAGACCTGCTGAAGAGTGTATTCAG 1109
Db      1100  TTGTGCTCAGGTTTTCATCCAAATGATGAGACCTGCTGAAGAGTGTATTCAG 1159
Qy      1110  TTAGGGTAAAGATTCATCAACACCTAAAGCTCAGGACCTTATACAGCAAAAGG 1169
Db      1160  TCAAGGTTAAAGATTCATCAACACCTAAAGCTCAGGACCTTATACAGCAAAAGG 1219
Qy      1170  ACATGATTAANGTGTGTGTACTACTAGCATAGTGTCAAGTCAATGTCAATGT 1229
Db      1220  ATATGATTA-----GACTGTTTGAATAGTATTTGAATCAATTAATTAACAATGC 1273
Qy      1230  TGGACATTTGTGAGCATGCGATGCTTGGTCAATGTTTCTGGCAAGGTTATCAGTA 1289
Db      1274  TTGATATATGAGCAACGATGCTTGGCAATATGTTTCTGCTAAGTCT----- 1327
Qy      1290  TTTGCTATATGGAAGATTAATATATCTGTGATGTGTGCTACCAAGTAATAGT 1349
Db      1328  TTTCAATTTTGAAGATTTGGGCAATCTGTTGATGTGTGCTACCAAGTAATAGT 1387
Qy      1350  TTTCTGTGCTACTTATCATCAAGATCTGTGATGAGGAAGTCAATACAGCAAGT 1409

```

```

Db      1388  TATCTTGACCTCGATTCATCAAACTCTGAGTGTGATTAATCAACAG-----G 1441
Qy      1410  AACTTGAACATGATGATGTAAGAGCTTGAGAAAATAGCAATTTGTGTACTTACAGCA 1469
Db      1442  AACTTGAATGATGATGTAAGAGCTTGAGAAAATAGCAATTTGTGTACTTACAGCA 1501
Qy      1470  GGGGATTAATTTCACTTATGCGAAATGTGAGCAATCTGCTGTATCTAGTAAGAAAGG 1529
Db      1502  GATGATTAATTTGTTAATTTGAATGTGAGAGGCTCATCTTGTATCTTGAAGAGCTT 1561
Qy      1530  GACGTGTCTGAGAAAAGTGGGTTAATGTTCAATGATCTCGAAGAGAGCGTCAAG 1589
Db      1562  TCAAGTTTACGTAACAATAGAGTTAATGTTCAATGATCTCACAGAGGAGCATCAAG 1621
Qy      1590  TTAACATGTGCTGATATGTCATGATAGAGGATGCAAGGACATCTGTGAAGCCCTTATC 1649
Db      1622  TGAACATTTCTGTTGTGCTCATGACAGTGAAGCAAGTGTCTCAAGCCCTCACT 1681
Qy      1650  AGCGTTCTTGAAGAGATGCTCTATCAAGTCAAG 1688
Db      1682  CAGCATCTTTGAGAGTGTCTTCTGCCAGAGTCAATG 1720

RESULT 8
AF135862
LOCUS
DEFINITION
  AF135862 1695 bp mRNA linear PLN 01-JUL-1999
  Glycine max precursor monofunctional aspartokinase mRNA, complete
  cde
ACCESSION
  AF135862
VERSION
  AF135862.1 GI:5305739
SOURCE
  ORGANISM
    Glycine max (soybean)
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids I; Fabiales; Fabaceae; Papilionoideae; Phaseoleae;
    Glycine.
  1 (bases 1 to 1695)
  Esau, B.D., Frankard, V., Jacobs, M. and Matthews, B.F.
  Isolation and characterization of a cDNA clone encoding a
  monofunctional aspartokinase
  Unpublished
  2 (bases 1 to 1695)
  Esau, B.D. and Matthews, B.F.
  Direct Submission
  Submitted (18-MAR-1999) SARR, USDA, 10300 Baltimore Ave.,
  Beltsville, MD 20705, USA
FEATURES
  source
    1..1695
    /organism="Glycine max"
    /mol_type="mRNA"
    /cultivar="Century"
    /db_xref="taxon:3847"
    1..1695
    /codon_start=1
    /product="precursor monofunctional aspartokinase"
    /protein_id="F4AD4176.1"
    /db_xref="GI:5305740"
    /translation="MASHALQQLQGVQGLAVSMYSVRSLHCKSCQIGPAALGAPVCAAR
    /RWMGNRAVFTVTCCKASTSDVLEKATENGWVSSRGESFTVMFGSSVASADRMK
    /EVALITLSPERPIVVLAMGKTNKLILAEKRVSCGVIVSSIEE.CFKDLHLR
    /TVDQLGVDSVIEKHLIELEQLKJAMKEKLTQDYLVLFVFGCMSTRIFAALINK
    /IGVARYDAFEIGFTTDDFTNADILLEATYPAVAKRLHGDMLSDPALAIVPGFKA
    /RKSQAVTLTGRGSDLLATTTIGALALPEILOYKVDGLTCDPNIPKAEVPLVLSI
    /DEAEELAPGAIVLHPSMRPAREDIPRVKNSVNPRAFGTLITKADMSAVLTLSI
    /VLRKNTVMDIATRMIGQYGFARVFSFEELGISVDYVATSEVSLTIDPSLNS
    /RELIOQASEIDHVELEKIAVNLQNSIISLIGNVRSLLIERLSRVLTGVT
    /VQMSQASRVNISLVVNDSEAEQCVRALHSAFPESELESEMDYKNQNGSVDELIS"

ORIGIN
  Query Match 37.4%; Score 742.2; DB 8; Length 1695;
  Best Local Similarity 72.4%; Pred. No. 1.9e-169;
  Matches 1008; Conservative 0; Mismatches 369; Indels 15; Gaps 3;

```

QY 284 GTGATGAAGTTCGGGGGCTCTGCTGTGTGGCCGCGAGATGGCTGAAGTGGCCGC 343
 Db 259 GTATGAAAGTTTGTGGCTCTGTGGCTTCTGTATGATGATGAAAGGTGGCTACC 318
 QY 344 CTGATCCTGAGTTCCTCCGAGAGCCGCCGTGCTGCTCTCTCCGATGAGGAAAC 403
 Db 319 CTATATATGAGTTTCCCGAGGAGAGCCGCTATGTGTTCTCTCTGCTATGGGAAAC 378
 QY 404 ACCAACCACTTCTCTCTGTGTGAGAGAGGAGTGGGTGGAGTTATCCATTTCT 463
 Db 379 ACAACCAAGCTTTGTGCTGTGAGAGAGAGCTGTGAGTTGGTGTATCAATGTATCA 438
 QY 464 GAATCGAAGATGGAATATGTCAAAGCCTATCATATCAAGCGGTGGATGAATTTGA 523
 Db 439 AGATATGAGAGCTTGTGCTTATTAAGACCTGATCTAAGAGCTGTGGATCAGCTTGT 498
 QY 524 CTTCGAAGATCTGTATACAA--GCCTTATGAACTGGAGCACTATGAAAGTATC 580
 Db 499 GTGAGCGATCTGTTATTTCAAGCATCTAAGAAATTTGAGCACTTCTGAAGGGATA 558
 QY 581 GCTATGATGAAGAGCTGACGCTGAGACGAGTACTACCTTGTTCATTTTGAAGATGC 640
 Db 559 GCTATGATGAAGATGACTAAGAGACTCAAGACTATTTAGTCTCTTGGAGATGC 618
 QY 641 ATGTCCACAGATTTTCTGCTTATTTGAACAAATTCGTGTCAAGGACGCGATAT 700
 Db 619 ATGTGACAGAGATCTTGTCTGTATCTTAAATTAAGTGTGTCAAGGCTCGCCAAAT 678
 QY 701 GAGCATTTGATTTGTTTCATTAACAATGATGATTTGATATGCGGATATCTTAA 760
 Db 679 GATGCAATTTGAGATTTGTTTATTAACAATGATGATTTGATATGCGGATATTTGAA 738
 QY 761 GCAACTATCTCTGCTGTGGAAGAGACTTCATGAGGATGAGTACAGATTCACGAT 820
 Db 739 GCAACTATCTCTGCTGTGGAAGAGACTTCATGAGGATGAGTATGCTCTCGATCTGCA 798
 QY 821 CCGTGTGATCTGAGGCTCTGCTGGAAGAGGCTGAAATCTGAGTGTATCACTTTAGC 880
 Db 799 GCAATTTGATCTGCTCTGCTGGAAGAGGCTGAAATCTGAGTGTATCACTTTAGC 858
 QY 881 CGAGTGTATGATGATCTGCTGCTGGAAGAGGCTGAAATCTGAGTGTATCACTTTAGC 940
 Db 859 AGAGGGGCGATGATTTGAGGCTGCTGGAAGAGGCTGAAATCTGAGTGTATCACTTTAGC 918
 QY 941 CAGATGGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1000
 Db 919 CAGATGGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 978
 QY 1001 AAGACTGTTCCATTAACATTTGAGAGGCGACAGACTTGTATTTGTGCTGAG 1060
 Db 979 GAACCTGTTCCATTAACATTTGAGAGGCGCGAGAGCTGATCTTTGTGCTGAG 1038
 QY 1061 GTTTTGATTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1120
 Db 1039 GTTTTGATTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098
 QY 1121 AATTTCATCAACCTTAAGCTCCAGGACCTTATTAACAGCAAGAGAGCATGGATTAAN 1180
 Db 1099 AATTTCATCAACCTTAAGCTCCAGGACCTTATTAACAGCAAGAGAGCATGGATTAAN 1157
 QY 1181 GGTGTGTTGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1240
 Db 1158 -----GGAGTATTAACAGCATTTTGAAGAGTGAATGATGATGATGATGATGAT 1212
 QY 1241 AGCACTCGATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1300
 Db 1213 AGCACTCGATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266
 QY 1301 GAAGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1360
 Db 1267 GAAGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326

QY 1361 CTTGATCATCAAGATCTGAGTGAAGCACTGATATACAGAGCAAGTGAACCTTGACCAT 1420
 Db 1327 CTGATCATCAAGATCTGAGTGAAGCACTGATATACAGAGCAAGTGAACCTTGACCAT 1386
 QY 1421 GTATGGAAGCTTGAAGAAATGAACATTTGCTCTACTTCAACAGAGGCGATTAAT 1480
 Db 1387 GTTGTGAAGAACTGAGAAATCTGCTGTGATCTCTTCAAGAAATGATTCATATC 1446
 QY 1481 TCACCTATCGAAGATGAGAGCAATGCTCTGATCTACTAGAAAGACGCGAGCTGTG 1540
 Db 1447 TCTCTATTTGAAGATTTCAAGATCATCAATATTTGAGAGCTTTCCCGTGTCTT 1506
 QY 1541 AGGAAAGTGGGTTAATGTTTCAATGATCTTCAAGAGGCTGCAAGGCTTAACATGTCG 1600
 Db 1507 CGAAGCTTGGGCTGATCTGTGCAATGATCTCCAGAGGCTGATCTAAGTGAACATCTCA 1566
 QY 1601 CTGATGCTCATGATGAGATGAGCAAGGACCTCGTAGAAGCCTTCATCAGGCGTCTTT 1660
 Db 1567 TTGCTTTAATGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1626
 QY 1661 GAAGAGATGTC 1672
 Db 1627 GAGAGTGAAGCTC 1638

RESULT 9
 AX507049
 LOCUS AX507049 1710 bp DNA linear PAT 27-SEP-2002
 DEFINITION Sequence 1744 from Patent WO0216655.
 ACCESSION AX507049
 VERSION AX507049.1 GI:23388286
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
 TITLE Stress-regulated genes of plants, transgenic plants containing
 same, and methods of use
 JOURNAL Patent: WO 0216655-A 1744 28-FEB-2002;
 The Scripps Research Institute (US) ; Syngenta Participations AG
 (CH)

FEATURES
 source location/Qualifiers
 1..1710
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"

ORIGIN
 Query Match 35.6%; Score 707.6; DB 6; Length 1710;
 Best Local Similarity 70.5%; Pred. No. 4.9e-161;
 Matches 1008; Conservative 0; Mismatches 401; Indels 21; Gaps 4;

QY 254 ATGAGAGATTTGGGAGATCACTGACGCTGTGATGAAGTTGGGGGCTCTGCTGTCG 313
 Db 235 ACCGAGATGATGAGAAAGGATATCAGTGCCTGATGAAGTTGGTGATCTTCGTTGGCG 294
 QY 314 TGGCGCGAGAGATGCTGAGGTGGCGGCGCTCATCTCGAGCTTCCCGAGAGAGCGCC 373
 Db 295 TCACTGAGAGATGAAGAGAGTTGCTGATTTGATTTTGAATTTCCGAGAGAGTCCC 354
 QY 374 GTGCTGTTCTCTGCTGATGAGGAAACCAACCAACCTTCTCTGCTGAGAGAG 433
 Db 355 GTCATGTTCTCTGCTATGAGGAAACCTCAACCAATCTTCTGCTTGGGAGAGAG 414
 QY 434 GCAATGAGGTGTGAGCTATCCATGTTCTGAATTCGAAGAGTGAATATGCTCAAGC 493
 Db 415 GCGGTAGTTGGTGTCTTCAATGATCTGAAGTGAAGGTTGAGACATTAATAAGAA 474
 QY 494 CTACATATCAAGCGGTGAATGAACTTGACT--TCCAGNATCTGTAAATACAGGCTT 550

Db 475 TTGCATATCAGAGCGGTGAAAAGAGCTCAACATTGATCCTCTGTATTTTGGACCTATTTG 534
 Qy 551 TATGAACTGAGCAACATATGAAAGGTATCGCTATGATGAAAGAGCTGACGCTAGAGCC 610
 Db 555 GAGGAACTGAGCAACCTCTGAAAGGCAATGCCATGATGAAAGAAATGACACTTGCAGCC 594
 Qy 611 AGTGAATACCTGTTCTTATTTGAGAAATGATGTCACACAGATTTTTTCTGCTTATTTG 670
 Db 595 AGAGATTAATCTAGTCTCTTTTGGAGAGGTTGTCTCAAGAGATTTTTTGTCTTATCTT 654
 Qy 671 AACAAATTCGTGTCAGAGCAAGGAGATGAGCATTTGATATTTGTTTCAATTAACAT 730
 Db 655 AATCAATCGGTGTCAGAGCAAGGCAATGATGATGATTTGAAATTTGTTTCAATTAACG 714
 Qy 731 GATGAATTTGGTAAATGCGGATATCTTAAAGCAACCTATCTCTGTTGGAGAGACTT 790
 Db 715 GATGATTCACAAATGCGGATATCTGAGAGCAATTCACAGCTGTGGCAAGATTTA 774
 Qy 791 CATGGGAGCTGGATACAGATACCGCATCTGTTGTTCTGGGTTCTTGGAGAGGCG 850
 Db 775 TATGATGATTTGGATGCAATCTCTGCTGTTCTTATTTGAACAGGTTTCTTGGAGAGGCT 834
 Qy 851 TGGAAATCTGCTGCTGATCTTACTTAAAGCCGAGTGTAGTGACTTGACTGCTACACC 910
 Db 835 TGGAAATCTGCTGCTGATCTTACTTAAAGCCGAGTGTAGTGACTTGACTGCTACACC 894
 Qy 911 ATTGTTAAAGCTTTGGAGCTGAGAAATTCAGGTATGAAAGATGTTGATGTTGATCTT 970
 Db 895 ATTGTTAAAGCTTTGGAGCTGAGAAATTCAGGTATGAAAGATGTTGATGTTGATCTT 954
 Qy 971 ACTTGTATCCAAATATCTTACCAATGCAAGAGCTGTCTCATCTTAACATTTGAAGAG 1030
 Db 955 ACATGTGACCTCACTATATATAGAGGCTACACAGATCAATCTGACATTTGACGAA 1014
 Qy 1031 GGCACAGAACTTGTCTTATTTTGGTCTCAGTTTTCATCCAAATGATGAGAACTGCT 1090
 Db 1015 GCAGCGGAGCTGATTTATTTTGGTGCACAGGCTCTTGACCCACAGTCAAGAACACAGCA 1074
 Qy 1091 AGAGAAAGTGAATATTCAGTTAGGCTTAAAGATCATACAACTTAAAGCTCCAGGACCC 1150
 Db 1075 AGAGAGGCTGATTTCTGTTAGGCTTAAATTTCTTAAACCTTAAGCTCTCTGAGAAC 1134
 Qy 1151 CTATATTCAGAAACAAAGACATGATTAAGCTGTGTTGATCTTAATGACATAGTCTC 1210
 Db 1135 ATCATCACTTAAACAAAGACATGACCAAGAG-----TATCTTAAAGACATTTGTTG 1188
 Qy 1211 AAGTCAATGTCACTATGTTGAGATTTGAGCACTCGAGTGTGCTGATGTTGTTT 1270
 Db 1189 AAGCTATGTGACCACTGCTGATTAATGCAAGCACCGAATGCTTGTCTCAAGTTGGCTTT 1248
 Qy 1271 CTGCGAAGGATTCAGTATTTGCTATATGAGATCTTATGATATCTGTGATTTGTTT 1330
 Db 1249 CTGCGAAGGATTTCTCATAT-----TTGAGAGAGCTTGGCAATTTCCGATGATGTTGTT 1302
 Qy 1331 GCTACCAAGTGAAGTATGTTCTGCTGCACTTATGATCCAAAGATCTGAGATGAGGAA 1390
 Db 1303 GCCACTATGAAATGATATATCTCTGACACTGATCTCTCAAACTTTGAGACAGAGAA 1362
 Qy 1391 CTGATACAGCAGGCAAGTGAATTGACCATGTAGTTGAAGAGCTTGAAGAAATGACATTT 1450
 Db 1363 CTGATTCACAG-----GAGCTGATCATGTAGTTGAAGAACTGAGAAATTTGACAGTT 1416
 Qy 1451 GTTGGCTACTTTCAGCAGAGAGGCGATATTTCACTTAATGCGAAATGAGCAATGCTCT 1510
 Db 1417 GTGATCTTTTAAAGAAAGCAATCATCTCTTAATTTGGAAATGTTCAACATTTCTCTCC 1476
 Qy 1511 CTGATACAGAAAGACGAGAGCTGTGCTGAGAGAAAGTGGGCTTAATGTTCAAGATGATC 1570
 Db 1477 CTGATTTTAAAGAGGCGCTTCAATGTTCTTATATCAAAAGTGTCAATGTCCAGATGATA 1536
 Qy 1571 TCGCAAGAGAGCTCAAAAGTGAATGATGCTGATGATGATGATGATGATGATGATGATGAT 1630
 Db 1537 TCACAAAGAGATCCAAAGTAAACATTTCTTTATATGTAACGAAAGCTGAGAGGT 1596

Qy 1631 CTCGTAGAGAGCCCTTATCATAGGCGCTTCTTGAAGACGATGCTCATACACA 1680
 Db 1597 TGTGTTAGAGCTCTTCAAAATCTTCTTGAGAGCGGTGACCTCTCAGA 1646
 RESULT 10
 BT000493
 LOCUS
 DEFINITION
 Arabidopsis thaliana aspartate kinase (At5g13280/T31B5_100) mRNA, complete cde.
 ACCESSION
 BT000493
 VERSION
 KEYWORDS
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1710)
 REFERENCE
 AUTHORS
 Cheuk, R., Chen, H., Kim, C.J., Shim, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Becker, J.R.
 Arabidopsis ORF clones
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Cheuk, R., Chen, H., Kim, C.J., Shim, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Becker, J.R.
 Direct Submission
 Submitted (25-SEP-2002) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 COMMENT
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFL CDNA (RFL CDNA : 'RIKEN Arabidopsis Full-length CDNA') : Seki, M., Narusaka, M., Ishida, J., Hayashizaki, Y., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 The Salk, Stanford, PEGC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RFL CDNA: Cheuk, R., Chen, H., Kim, C.J., Shim, P., Bower, J., Chan, M.M., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Oodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Tortum, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W., Theologis, A., and Becker, J.R.
 Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Becker, J.R. (SSP/Salk) contributed equally to this work as PIs.
 Location/Qualifiers
 1..1710
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="U4700"
 /ecotype="Columbia"
 /note="This clone is in PUNI 51"
 1..1710
 /note="aspartate kinase"

CDS


```

/codon_start=1
/product="A59:3280/T31B_100"
/protein_id="A518062.1"
/db_xref="GI:23308185"
/translation="MAATRVCCSHNAFRLPLTRHNSPLPISLRVDPFLTKL
SLPIDGSSIRKVSQSRNIVRAVLEKTEATTEVEDEKITCVMKFGSSVSAER
MEVADLITPEPESPVIVLSMGKTNLTLAGEKAVSCVNSAIEELSLIKELH
IRVKEINIDPSVILTYLEELBOLKIGLAMKELTLERDVLVRSGLSTRIFAAVL
NTRVRAVOYPAFEIGPTTDDPTDDELTATVPVAKRPLDMDMPAPAPITVTEG
KMKRTGAVTTLRGSSDLTATTIGKALIKELIQVKKVVDGLTIDPTTYKRAIPVPL
TFDEAEIAYFPAQVLHPQSMRPAEBSIPIRVKNSINPAPGTTIKTIDMKSLIT
SIVLRNVTMLDIATRMIGQVGLPAKVFSLPEELGISVDAVASEVISLTLPSGL
WSRELIOQLDHLVBELEKIAVNLKRAIISLIGNVSHSLLEBAFVLYTKGVN
VOMISQASKVNIISPIVNEAEAGCVQLHKSFPESGSLSELILQPLNGNSPRTLO
VEN"

```

ORIGIN

```

Query Match      35.6%; Score 707.6; DB 8; Length 1710;
Best Local Similarity 70.5%; Pred. No. 4,9e-161;
Matches 1008; Conservative 0; Mismatches 401; Indels 21; Gaps 4;

QY 254 ATGGAGGGATGGGGGATCAGCTCAGCGGTGATGAAGTCCGGGGGCTCTCGGTGCG 313
DB 235 ACGGAGGTGATGAGAGGGATATCAGTGTGATGAGAGTTGGTGAATCTCGGTGCG 294
QY 314 TCGGCGCGAGGATGCGTGAAGTGGCGCGCTCATCTGACGTTCCCGAGAGCGCCC 373
DB 295 TCAGCTGAGAGATGAGAGAGTGTCTGATTTGATTTGACCTTTCCGGAAGAAAGTCCC 354
QY 374 GTGCGTCTCTCTCGCCATGGGGAAAACCAACAACCTTCTCTGCTGAGAGAG 433
DB 355 GTCATGTTCTCTCTGCTAAGGGGAAAACCAACAACCTTCTCTGCTGAGAGAG 414
QY 434 GCGATGAGGTGAGATTATCCATGTTTCTGAATCGAAGTGAATATGTCMAAAGC 493
DB 415 GCGGTTAGTTGGTGGTTCCTAATGATCTGAGATTGAGAGATTGAGCATTAAGAA 474
QY 494 CTACATATCAAGCGGTGATGAATCTTGACT--TCCAGNATCTGTAATCAAGCTT 550
DB 475 TTGCATATGAGACGGTGAAGAGCTCAACATGATCCCTGTTATTTGACCTATTTG 534
QY 551 TATGAACGTGAGCAACTATTGAAAGTATCGCTATGATGAAAGAGCTGACCCATGAGC 610
DB 535 GAGGAATCGAGCAACTCTCTGAAAGCAATGCGCATGAGGAAGATTGACACTCGAACC 594
QY 611 AGTACTACCTGTTTCATTGTAAGATGATGCAACAGAGATTTTTCGCTTATTTG 670
DB 595 AGGATTTACTTAGTCTCTTTTGAGAGTGTGTTGTCACAGAGATTTTTCGCTTATTT 654
QY 671 AACAAATTCGATGCAAGGACGCGAGATGACGATTTGATGATGTTGTTTATTAACAT 730
DB 655 AATACATCGGTGTCAAGGACGCGCAATGATGATTTGAAATTTGTTTATTAACAAG 714
QY 731 GATGAATTTGTAATGCGGATATCTTGAAGCAACTTATCTGCTGTTGGAAGAAGCTT 790
DB 715 GATGATTTCACAAAATGGGATATCTCGAGGCAACTTATCCAGCTGTTGCCAAGAATTA 774
QY 791 CAGGGGACGTGATTCAGATTCAGAGTACCTGTTGTAATCGGGTCTCTGGGAAGGCG 850
DB 775 TATGATGATTTGATGATATCTCTGCTGTTCTTATTTGAACAGGTTCTTTGGGAAGG 834
QY 851 TGGAAATCTGTGCTGTAACTACTTTAGCCGAGGTGATGATGATGATGATGATGATGAT 910
DB 835 TGGAAATCTGTGCTGTAACTACTTTAGCCGAGGTGATGATGATGATGATGATGATGAT 894
QY 911 ATTGTAAGCCTTGGGATGAGAGAAATTCAGATGATGATGATGATGATGATGATGATGAT 970
DB 895 ATTGTAAGCCTTGGGATGAGAGAAATTCAGATGATGATGATGATGATGATGATGATGAT 954
QY 971 ACTTGTATCAAAATATCTTACCAACATGCAAGAGCTGTTCCATCTTAACATTTGAAGAG 1030
DB 955 AACTGTGACCCCTATATATATAGGAGCTTACACAGATGATCCATATCTGACATTTGACGAA 1014

```

```

QY 1031 GCCACAGAACTGTGATTTTGTGTGCTCAGGTTTGCATCCACATGATGATGATGATGAT 1090
DB 1015 GAGCGGAGCTAGCTTATTTTGTGTGCAAGGTCTTGACCCACAGTCAATGACCAAGCA 1074
QY 1091 AGAGAGGTGATATTTCCAGTTAGGTTAAGATTCATACCAACCTTAAGCTCCAGGACCC 1150
DB 1075 AGAGAGGTGATATTTCCAGTTAGGTTAAGATTCATACCAACCTTAAGGCTCCGGAAC 1134
QY 1151 CTATTACCAACAAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1210
DB 1135 ATCATCATTAACCAAGACATACCAAGAG-----TATTTACAGGATGTTGTTCTG 1188
QY 1211 AAGTCAAAATGCTCATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1270
DB 1189 AAGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1248
QY 1271 CTGGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1330
DB 1249 CTGGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1302
QY 1331 GCTACAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1390
DB 1303 GCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1362
QY 1391 CTGATACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1450
DB 1363 CTGATACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1416
QY 1451 GTTGTCTTACTTCAAGACAGAGGCGATATTTCACTTATCGAAATGAGAGCATGCTCT 1510
DB 1417 GTGATCTTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1476
QY 1511 CTGATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1570
DB 1477 CTGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1536
QY 1571 TCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1630
DB 1537 TCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1596
QY 1631 CTCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB 1597 TGTGTTCAAGGCTCTTCACAAAATCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1646

```

```

RESULT 11
LOCUS AY057674 1971 bp mRNA linear PLN 22-OCT-2001
DEFINITION Arabidopsis thaliana AT5g13280/T31B_100 mRNA, complete cds.
ACCESSION AY057674
VERSION AY057674.1 GI:16323140
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1971)
REFERENCE 1
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,J., Kamuya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinzaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE Arabidopsis cDNA clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1971)
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,J., Kamuya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,

```

TITLE

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinzaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

COMMENT

Direct Submission
Submitted (24-SEP-2001) Salk Institute Genomic Analysis Laboratory (Signal), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFI CDNAs (RAFI CDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFI CDNAs: Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shim, P., Ban, J., Bowser, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

1. 1971
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="RAF09-49-Q22(R14700)"
/ecotype="Columbia"
1. 1788
79. 1788
/note="aspartate kinase"
/codon_start=1
/product="AT5G13280/T31B5_100"
/protein_id="At1G3305.1"
/db_xref="gi:16323141"
/translation="MAATVRRCCHSNAAFTRLPLTRHNSPTLPLISINRYDPTLKL
SLPFDGSSIRKVSGRNIVRAVLEKTEATTEVDKQITCWMFGSSVASKR
MKEVADLLITPEPESPVILSAMKGTINLLAGEKAVSCVNSASIELSLIKELH
IRTVELNIDPEVILVLELEBOLKIGAMKELTLRLDVLVSGELSTR.FAAVL
NTIGKAROYDAFEIGFTTDDPTNGDILEATYPAVKRLYDDMDHPAVPIVTFGL
KGMKGAATTTGGGSDLTATTTGKALKEIOWMKVDVLTCDPITVYRAAPVYLT
TFDEAELAYRGAQVYTHPSRPARBEGIPYRVKNSVNPAPGIIIRKTDKSKILT
STLVKRVATMDIATRVLGOVGLAKVFS.FEELGVSVDVATSEVSLITLDPKSL
WSRELIOQLDHVLELEKIAVNLKRAIISLIGVQHSLLLEAFALVLTQGN
VOMISOGASXVNIPIVNEAEBCVQALHRSFPESDLELLIQPLGNSVPTLQ
VEN"
1789. 1971

ORIGIN

3' UTR

Query Match 35.6%; Score 707.6; DB 8; Length 1971;
Best Local Similarity 70.5%; Pred. No. 5e-161;
Matches 1008; Conservative 0; Mismatches 401; Indels 21; Gaps 4;

QY 254 ATGAGAGGATGGGGATCAGCTCAGCGTGTGATGAAGTTCGGGGGCTCTCGGTG 313
DB 313 ACGGAGGTGATGAAAGGATATCAGTGCCTGATGAAGCTTTGGTGATCTTCG 372
QY 314 TCGCCGCGAGATGAGCTGAGTGCCTGATCTGACGTTCCCGAGAGAGCCCC 373
DB 373 TCGAGTGAAGATGAAGAGAGTGTGATTTGATTTTCCGGAAGAAAGTCCC 432
QY 374 GTGGTGTCTCTCTGCTGATGAGGAAACACCAACAACCTTCTCTCTGAGAGAG 433
DB 433 GTGATGTCTCTCTCTGATGAGGAAACACCAACAACCTTCTCTCTGAGAGAG 492
QY 434 GCAGTAGGCTGAGATATTCATGTTTCTGAATCAAGAGTGAATATGCTCAAAAGC 493

DB 493 GCGGTATGTTGTGTGTTCTTAATGATCGATGAGATGAGAGTTGACCTTAAGAA 552
QY 494 CTACATATCAAGACGGTGAATGAACTTGACT---TCCAAAGNATCTGTATACAAAGCTT 550
DB 553 TTGATATACAGACGGTGAAGAGCTCAACATGATCCCTCGTTATTTTACCAATTGG 612
QY 551 TATGATCTGAGCAACTTATGAAAGGTATCGTATATATGAAAGAGCTGACCTTGAGAC 610
DB 613 GAGAACTGAGCAACTCCGAAAGGATGCGCATATGAAAGATTTGACCTTGAGAAC 672
QY 611 AGTACATACCTTTGTTGATTTGGAGATGATGCAATGCAAGATTTTCTGCTTATTTG 670
DB 673 AGAGTTTCTTAGTCTCTTTTGGAGAGTGTCTTCAAGAGATTTTTCCTGCTTACTT 732
QY 671 AACAAATTCGTGTCAAGGACGCGCATATGACGATTTGATTTGTTTCAATACACT 730
DB 733 AATCAATCGGTGTCAAGGACGCGCAATATGATGATTTGAAATTTGTTTCAATACAG 792
QY 731 GATGATTTGTTATGCGGATATCTTATGAGCAACCTATCGCTGTGCGAAGACTT 790
DB 793 GATGATTTACAAATGGGAGATTCCTGAGAGCACTTATTCAGCTTGTCCAAAGATTA 852
QY 791 CATGGGACTGATACAGATCCAGCATACCTGTTGTTACTGGGTTCTTGGAAGAGGC 850
DB 853 TATGATGATGATGATGATGATCTGCTGTTCTATTTGAAAGAGTTCTTGGAAGAGGC 912
QY 851 TGGAAATTCGTGTGTTAACTTCTTATGCGCGAGGTGTATGATCTGATCTCAAC 910
DB 913 TGGAAATTCGTGTGTTAACTTCTTATGCGCGAGGTGTATGATCTGATCTCAAC 972
QY 911 ATTGGTAAAGCTTGGAGCTGAGAGAAATTCAGGTATGAAAGATTTGATGTTACTT 970
DB 973 ATTGGTAAAGCTTGGAGCTTGGAGAAATTCAGGTATGAAAGATTTGATGTTACTT 1032
QY 971 ACTGTGATCCAAATATCTACCAATGCAATGCAAAAGCTTTCATCTTAATTTGAAG 1030
DB 1033 AATGTGATCCCTATATATATATGACGAGCTACCAAGTATGATGATTTGACGA 1092
QY 1031 GCAAGAGATTCGTGTTATTTTGTGCTGAGTTTGTGATGCAATGATGAGACCTGCT 1090
DB 1093 GCAAGAGATTCGTGTTATTTTGTGCTGAGTTTGTGATGCAATGATGAGACCTGCT 1152
QY 1091 AGAGAGGTGATTTCCAGTTAGGTTAAGATTTCAATCAACCTTAAAGTCCAGGAC 1150
DB 1153 AGAGAGGTGATTTCCAGTTAGGTTAAGATTTCAATCAACCTTAAAGTCCAGGAC 1212
QY 1151 CTATTTACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1210
DB 1213 ATCATCTAATAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266
QY 1211 AAGTCAATGCTACTATGTGAGATTTGATGATGATGATGATGATGATGATGATGAT 1270
DB 1267 AAGTCAATGCTACTATGTGAGATTTGATGATGATGATGATGATGATGATGATGAT 1326
QY 1271 CTGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1330
DB 1327 CTGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1331 GCTACAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1390
DB 1381 GCTACAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1391 CTGATAG 1450
DB 1441 CTGATAG 1494
QY 1451 GTTGTCTACTTACAG 1510
DB 1495 GTTGTCTACTTACAG 1554
QY 1511 CTGATAG 1570

Db 1555 CTGATTTTAGAGAGGGCGTTTCATGTTCTTTATACCAAGTGTCAATGTCCAGATGATA 1614

Qy 1571 TCGCAAGGAGCGTCAAGGTATACATGTGCGATAGTCCATGATAGCGATCAAGGCA 1630

Db 1615 TCAACAGGAGCATCCAGGTAAACATTTCTTATATAGTAAAGCAAGCTGAAGGT 1674

Qy 1631 CTCGTGAAGCCCTTCATCAGCGCTTTCTTTGAAGCATGTCTTATCA 1680

Db 1675 TGTGTTCAGGCTCTTCAAAATCTTCTTCGAGAGCGGTGACCTCTCAGA 1724

RESULT 12

ATAKLYSI 1980 bp mRNA linear PLN 02-JUL-1997

LOCUS A.thaliana mRNA for aspartate kinase.

DEFINITION X98873

ACCESSION X98873

VERSION X98873.1 GI:2243115

KEYWORDS ak-1ys1 gene; aspartate kinase.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1

AUTHORS Frankard, V., Vauterin, M. and Jacobs, M.

TITLE Molecular characterization of an Arabidopsis thaliana cDNA coding for a monofunctional aspartate kinase

JOURNAL Plant Mol. Biol. 34 (2), 233-242 (1997)

MEDLINE 97351557

PUBMED 9207839

REFERENCE 2

AUTHORS Frankard, V.M.S.

TITLE Direct Submission

JOURNAL Submitted (02-JUL-1997) V.M.S. Frankard, Institute for Molecular Biology, Lab of Plant Genetics, Free University of Brussels.

Peptide: 65, 1640 Sint Genesius Rode, BELGIUM

Peptide: revised by [3]

3 (bases 1 to 1980)

REMARK

REFERENCE

AUTHORS Frankard, V.M.S.

TITLE Direct Submission

JOURNAL Submitted (02-JUL-1997) V.M.S. Frankard, Institute for Molecular Biology, Lab of Plant Genetics, Free University of Brussels.

Peptide: 65, 1640 Sint Genesius Rode, BELGIUM

Peptide: Related sequence: X99107.

Location/Qualifiers

1. 1980

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/note="ecotype Columbia"

1. 1980

/gene="ak-1ys1"

79. 1788

/gene="ak-1ys1"

/EC_number="2.7.2.4"

/codon_start=1

/product="aspartate kinase"

/protein_id="CAA67376.1"

/db_xref="GI:4376158"

/db_xref="GOA:O23152"

/db_xref="SPRPMBL:O23152"

/translation="MAATRVRCCHSNAAPTRLPTRHNSPTLPSLNRYDPFLKLSLPIGSSIRKVSIGSGSRNIVRAVLEKTKTEATVEDEKITCVKRGSSVASAKRKEVADLITPEESPVIIVLSAGKTTNNLLAGEKAVSCVNASSEIIBELIIEKHIRTVELINDPSVILTYLEBELBOLKGIAMKMKELTRRDIVSFGBELSRIIPAVLNTIGVARYDAFEIGFITTDFTNGDILEATYPAVARLIDMMHPPAVIVTGFGLKGMKTGAVTTLGRGSGDLTATIGKALKEIQWKVDVGLTCDPTLYKATVPVLTTPDEAEILAYFGAOVLHPOSNRPARBEGIPYKVSYPKAPGIIITKTRMTSILTSIVKXNTMTLIDIASTRMTCGOVPLAKYPTFEETGISVDVATSEVSLITLDPSTLWSRELIQELDHYVELEKIAIVANLIKGRALISLIGNVQHSILIEAPHLITKGINVQNIISGASKVNISFTVNEABEGCVQALHKSFFESGDLSELLIQPRGNSPVRTIQVBD"

79. 336

mat_peptide /gene="ak-1ys1"
337. 1785
/gene="ak-1ys1"
/product="aspartate kinase"
/EC_number="2.7.2.4"

misc_feature 339. 360
/gene="ak-1ys1"
/note="conserved aspartate kinase amino acid box, KRGG"
1970

polyA_site /gene="ak-1ys1"

ORIGIN

Query Match 35.6%; Score 706; DB 8; Length 1980;
Best Local Similarity 70.4%; Fred. No. 1.2e-160; Mismatches 1007; Conservative 0; Mismatches 402; Indels 21; Gaps 4;

254 ATGAGAGATTGGGGATCAGCTCAGCGTGTGATGAAGTTCGGGGGCTCTGGTTCG 313

Db 313 ACGAGAGTAAATGAAAGGATACACGTCGTGATGAAGTTGGATCTTCGGTGGC 372

Qy 314 TGGCCCGAGAGATGGCTGAGGTGGCCGCTCATCTGACGTTCCCGAGAGCGCCC 373

Db 373 TCAAGCTGAGATGAAGAAAGTGTGATTTGACTTTCCGGAAGAAAGTCCC 432

Qy 374 GTGATGTTCTCTCGCCATGAGGAAAAACCAACAACCTTCTCTTGGAGAGAAG 433

Db 433 GTCAATTTCTCTCTGCTATGGGAAAACCAACAATCTTCTTGGAGAGAAG 492

Qy 434 GCAATGAGGTGAGATTATCCATGTTCTTGAATTCGAAGATGAAATATGTCAAAAC 493

Db 493 GCGATTAGTGGTGATTCTTAATGCACTGAGATTGAGAGTGAACATTATTAAGAA 552

Qy 494 CTACATCAAGAGCGGTGATGAACCTTGACT--TCCAGNATCTGTAATACAGCTT 550

Db 553 TTGCATTCAGACGCTAAGAGCTCAACTTATCTCTGTATTTGACCTATTG 612

Qy 551 TATGAACCTGAGCAACTATTGAAGATTCGCTATGATGAAGAGCTGACGCTAGAAC 610

Db 613 GAGAACTGAGCAACTCTGAAAGCATTCGCAATGATGAAGATGACATTCGAAC 672

Qy 611 AGTACTACCTGTTTCTTATTTGGAGAAAGCATGTCACCGAGATTTTCTGCTATTG 670

Db 673 AAGATATCTTATGTCCTTTTGGAGAGTTTGTCTCAAGAGATTTTCTGCTATTG 732

Qy 671 AACAAATTCGTGTCAAGGACGCGAGTATGACCATTTGATATGTTTCATTCAACT 730

Db 733 AATACATGGGTCAAGGACGCGCATATGATGATTTGATATGTTTCATTCAACG 792

Qy 731 GATGAATTTGGTAATGCGGATATCTTAGAACCAACTATCTGTGTGGGAAGACTT 790

Db 793 GATGATTCACAAAGGGGATATCTTGAGGCAACTATCAAGTGTGCAAGAGATTA 852

Qy 791 CATGGGACCTGATACAGATTCACGCACTACCTGTTGACTGAGGTTCTTGGAGAGG 850

Db 853 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 912

Qy 851 TGAAGATCTGTCTGTATCACTATTAGCCGAGGTGTATGATCTGATCTCTCAAC 910

Db 913 TGAAGATCTGTCTGTATCACTATTAGCCGAGGTGTATGATCTGATCTCTCAAC 972

Qy 911 ATTGATGAAGCTTGGGCTGAGAGAAATTAAGTATGAAGAGTGTGATGAGTACTT 970

Db 973 ATTGATGAAGCTTGGGCTTGAAGAGATTAAGTATGAAGAGTGTGATGAGTACTT 1032

Qy 971 ACTTGATTCACAAATATCTACCAATGCAAGAGCTGTTCCATCTTAACATTTGAAG 1030

Db 1033 ACATGTGACCTACTATATATTAAGGAGCTCACACAGTACATATCTGACATTTGAG 1092

Qy 1031 GCAAGAACTGCTTATTTGGTGTGAGTGTGATGATGATGATGATGATGATGATGAT 1090

Db 1093 GCAAGAACTGCTTATTTGGTGTGAGTGTGATGATGATGATGATGATGATGATGAT 1152

Qy 1091 AGAAGAGGTGATATTCAGTTAGGTTAAGATTCATCAACCTTAAGCTCCAGGACCC 1150

1153 AGAGAGGGGAGATTCCTGTTAGGGTTAAAGTTCTTATACCTTAAGGCTCCCTGGAACT 1212
 QY 1151 CTTATTAACAGCAAGAGAGACATGATTAAGGTGCTGTTGACTAATCAGATAGTCTC 1210
 Db 1213 ATCATCTAATAAACAAGACATGACCAAGAG-----TATCTTAAGAGCATTTGTTCTG 1266
 QY 1211 AAGTCAATGCTACTATGTTGGACATTTGTGAGCACTGGATGCTTGGTCAATATGTTT 1270
 Db 1267 AAGCTAATGATGACCATGCTGATATACCAAGCAACCCGAATGCTTGGTCAAGTTGGCTTT 1326
 QY 1271 CTTGCAAGGTTATCAGATATTTGCTATATTTGAAGATCTATGATATATGCTGATGTTGTT 1330
 Db 1327 CTTGCAAGGTTATTCCTGATAT-----TTGAGAGCTTGGCATTTCCGTAGATGTTGTT 1380
 QY 1331 GCTACCACTGAAGTATGTTTCTGTGCTACTTGAATCCATCAAAAGATCTGGAGTAGGAA 1390
 Db 1381 GGCACCTAGTGAAGTCACTATCTCTGACACTGATCCTCAAACTTTGGAGCAGAGAA 1440
 QY 1391 CTGATACGACGAGCAAGTAACTTATACATGATGATGAAGCTTGAAGAAATAGCAAT 1450
 Db 1441 CTGATTCACAG-----GAGCTTATCATGTATGAAGACTGGAGAAATTTGGCAGTT 1494
 QY 1451 GTTCGCTACTCTCAGCAGAGGCGATTAATTTCACTTATCGGAAATGTGAGCAATGCTCT 1510
 Db 1495 GTGAATCTTCTAAAGGAAGAACATTCCTCTTAATTTGGGAATGTTCAACATTTCTCC 1554
 QY 1511 CTGATACCTGAAGAAAGCAGAGCTGTCTGAGGAAAGTGGGTTAATGTTCAATATATC 1570
 Db 1555 CTGATTTTGAAGAGGCGCTTTCATGTCTTATATACCAAGGTGTCAAATGTCATATATA 1614
 QY 1571 TCGCAAGAGCGCTCAAGGTTAACTGCTGATAGTCCATGATAGCATGACGACAAAGCA 1630
 Db 1615 TCACAAAGGACATCCAGGTTAAACATTTCTTATAGTAAAGAAAGCTGGAAGCTGGAAGT 1674
 QY 1631 CTGCTAGAAAGCCCTTCATCAGGCTTTCTTGAAGCATGTCTATCA 1680
 Db 1675 TGTGTCAGGCTCTTCAAAATCTTCTTCGAGAGCGGTGACCTCTCAGA 1724

RESULT 13
 LOCUS AY089207 1933 bp mRNA linear PLN 14-APR-2003
 DEFINITION Arabidopsis thaliana clone 94688 mRNA, complete sequence.
 ACCESSION AY089207
 VERSION AY089207.1 GI:21407981
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCES
 AUTHORS Haas, B.J., Volfovsky, N., Town, C.D., Troukhan, M., Alexandrov, N.,
 Feldmann, K.A., Flavell, R.B., White, O., and Salzberg, S.L.
 TITLE Full-length messenger RNA sequences greatly improve genome
 annotation
 JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
 MEDLINE 22088475
 PUBMED 12093376
 REFERENCE 2 (bases 1 to 1933)
 AUTHORS Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
 Feldmann, K.
 TITLE Full-length cDNA from Arabidopsis thaliana
 JOURNAL unpublished
 AUTHORS 3 (bases 1 to 1933)
 TITLE Direct Submission
 JOURNAL Submitted (11-MAR-2002) Ceres, Inc., 3007 Malibu Canyon Road,
 Malibu, CA 90265, USA
 COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made
 available at TIGR and Genbank. The following quality assessment of

this set was done by comparison with known proteins: two percent of
 the clones are estimated to be 5'-truncated; less than one percent
 are 3'-truncated; approximately two percent represent alternative
 splice variants, including unspliced introns and spliced exons; one
 percent may contain premature stop codons; five percent may have
 frame shifts in a coding region. A sequence is considered to be
 5'-truncated if it lacks the translation initiation start (ATG). A
 sequence is considered to be 3'-truncated if it lacks the
 C-terminal end of the encoded protein. Please note that these cDNA
 sequences are derived from the WS or Laer ecotypes and therefore
 may contain polymorphisms when compared to sequences from Col-0.
 Geneset carried out the library production and sequencing of the
 full-length clones. Ceres, Inc. carried out the clustering of the
 5' sequences, selection of clones, and sequence assembly.

FEATURES

source
 1..1933
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="94688"

ORIGIN

Query Match 35.0%; Score 695.6; DB 8; Length 1933;
 Best Local Similarity 70.4%; Pred. No. 4.2e-158;
 Matches 1007; Conservative 0; Mismatches 401; Indels 22; Gaps 5;
 254 ATGAGGAGATTGGGGATCAGCTCAGCGTGTGATGAAGTCCGGGGGCTCTGGTGTG 313
 Db 280 AGGAGAGTATGATGAAGAGGATACCTGCGTATGAAGTTGGTGTATTTGGTGGG 339
 QY 314 TCGCCCGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 373
 Db 340 TCAGCTGAGAGATGAAGAGAGTGTGATTTGATTTTGGCTTTCCGGAAGAAATGCC 399
 QY 374 GTTCGCTTCTCTCGCCATGGGAAACCCACCAACCTTCCTCTGCGAGAGAG 433
 Db 400 GTCAATTTCTCTCTGCTATGGGAAACCTACCAACATCTTGTCTGGGAGAGAG 459
 QY 434 GCATAGAGGTGTGAGTATTCATGTTTCTGAAATCGAAGAGTGAATATGTCAAAAGC 493
 Db 460 GCGGTATGTGGGTGTTCTAATGATCTGAGATGAGGAGTTGAGCATTAAGAGAA 519
 QY 494 CTACATACAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 550
 Db 520 TTGCATATCAGAGCGTAAAGAGCTCAACATGATCTCTGTTATTTGACCTATTTG 579
 QY 551 TATGAACCTGAGCACTATTGAAGGTATGCTATGATGAAGAGCTGAGCCTAGAGCC 610
 Db 580 GAGAACTGAGCAACTCTGAAAGCATTTGCCATGAT-AAAGAAATGACATTCGAAAC 638
 QY 611 ATGTCTACCTTTTCAATTGGGAATGATGTCACCAAGATTTTCTGCTATTTG 670
 Db 639 AAGATTTACTAGTCTCTTTTGAAGAGTGTGTCACAGAGATTTTGTGCTATCTT 698
 QY 671 AACAAATTCGTGTCAGAGCAGCAGATGATGACCATTTGATATTTGTTTCAATCACT 730
 Db 699 AATACATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 758
 QY 731 GATGAATTTGGTATGCGGATATCTTAGAAGCACTATCTGCTGTGGGAAGAGCTT 790
 Db 759 GATGATTTTCAAAAGGAGATATCTGAGGCACTTATGAGTGTGCAAGAGTTA 818
 QY 791 CATGGGAGCTGATCAGATTCAGAGCATCTGTTGATCTGAGGTTCTTGGAGAGGC 850
 Db 819 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 878
 QY 851 TGAATATCTGTGCTGTAATCTTTAGGCCGAGGTGTAGTACTGATCTCAACACC 910
 Db 879 TCGAAATCTGTGCGGTATCTAATCTTGGTGAAGGTGCGAGTATTTGAAGGCAACACA 938
 QY 911 ATTGATTAAGCTTGGAGCTGAGAGAAATTCAGGATGAGAAAGTGTGATGATGAT 970
 Db 939 ATTGATTAAGCTTGGTGTGATTAAGAGATTCAGGTTGGAAGATGTCATGATGTTCTA 998

QY 971 ACTTGATCCAAATATCTACCAATGCAAGACTGTTCCTACTTAACTTTGAAG 1030
 Db 999 ACATGTACCCTACTATATTAAGAGAGTACACAGTACCAATCTGACATTTGACGAA 1058
 QY 1031 GCCACAGACTTGTCTATTTGTGTCTGAGTTTTCATCCACATCGATGAGACTGCT 1090
 Db 1059 GCAGCCGACCTAGCTTATTTGTGTGACAGGCTTTCGACCCACAGTCAATGACACGCA 1118
 QY 1091 AGAAGAGTGATATTCAGATTGAGTTGAAGATTATTAACAACCTTAAGCTCCAGGACC 1150
 Db 1119 AGAGAGGTGAGATTCCTGTGTAGGGTTAAATTTCTTAAACCTTAAGGCTCCGGAAC 1178
 QY 1151 CTATATACCAAGACAAAGACATGATTAAGGTCTGTGTACTAATAGATAGTGTCTC 1210
 Db 1179 ATATCATCTTAAACAAAGACATGACCAAGAG-----TATTTACGAGCATTTGTCTG 1232
 QY 1211 AAGTCAATGTCACTATGTTGACATTTGACACTCGAGTCTTGTGATGATGTTT 1270
 Db 1233 AAAGTAAATGTGACATGCTGATATAGCAACACCCGATGCTTGTCAAGTTGGCTTT 1292
 QY 1271 CTGGCAAGGTATCAGATTTTGTCTATATTAAGATCTATGATATCTGTGATTTGTT 1330
 Db 1293 CTTCGAAAGGTATTCGATAT-----TTGAGAGCTTGGCATTTCCGTAGATTTGTT 1346
 QY 1331 GCTACCAAGTGAAGTATGTTTCTGTCTGCTGATTCATCAAGATCTGAGTAGGAA 1390
 Db 1347 GCCACTAGTAGAGTCAATCTCTGACATGATGATCTTCAAACTTTGAGCAGAGAA 1406
 QY 1391 CTGATACAGCAGCAGATGATCTGACATGATGATGATGATGATGATGATGATGAT 1450
 Db 1407 CTGATTCACACAG-----GAGCTGATCTGATGATGATGATGATGATGATGATGAT 1460
 QY 1451 GTTCGTCTACTTACAGCAGAGGAGGATTAATTCATCTTACGAAATGTCAGCAATCTGCT 1510
 Db 1461 GTGAACTCTCTAAAGAGAGAGAGATCATCTCTTAATTTGGGAATGTTCAACATTTCCCTC 1520
 QY 1511 CTGATCTAGAAAGACGAGAGCTGTGCTGAGAGAAAGTGGGTTAATGTTCAATGATC 1570
 Db 1521 CTGATTTTGAAGAGGAGGCTTTCATGTTCTTATTAACAAAGGTCAAGTCAATGATGATA 1580
 QY 1571 TCGCAGAGAGCGTCAAGGATTAACATGCTGATGATGATGATGATGATGATGATGATGAT 1630
 Db 1581 TCACAGAGAGCATCCAGGATTAACATTTCTTATTAAGTAAAGAGCTGAGAGCTGAGAGT 1640
 QY 1631 CTGTAGAAAGCCTTCTATCAGGCTTCTTGAAGCATGTCCTATCA 1680
 Db 1641 TGTGTCAGGCTCTTCAAAATCTTCTTCGAGAGGCTGATCTCTCAGA 1690

RESULT 14
 ATU62020 1923 bp mRNA linear PLN 15-JUL-1997
 LOCUS Arabidopsis thaliana lysine-sensitive aspartate kinase mRNA.
 DEFINITION complete cds.
 ACCESSION U62020.1 GI:2257742
 VERSION U62020
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1923)
 Tang, G., Zhu-Shimoni, J.X., Amir, R., Zchori, I.B. and Galili, G.
 Cloning and expression of an Arabidopsis thaliana cDNA encoding a
 monofunctional aspartate kinase homologous to the lysine-sensitive
 enzyme of Escherichia coli
 JOURNAL Plant Mol. Biol. 34 (2), 287-293 (1997)
 MEDLINE 97351562
 PUBMED 9207844
 REFERENCE 2 (bases 1 to 1923)
 TANG, G., ZHU, X.-Z. and GALILI, G.

TITLE Direct Submission
 JOURNAL Submitted (25-JUN-1996) Plant Genetics, The Weizmann Institute of
 Science, Rehovot 76100, Israel
 FEATURES
 source
 1. 1923
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 117..1751
 /note="similar to E. coli lysine-sensitive aspartate kinase"
 111, Swiss-Prot Accession Number P08660"
 CDS
 1..1923
 /product="lysine-sensitive aspartate kinase"
 /protein_id="AAB63104.1"
 /db_xref="gi:2257743"
 /translation="MASTOLYGVKTPGLATLSXRLFEPAKGFVTLPESSAVFRDV
 EHSRNLGLRVSCALRVLDLORKEPTDSGGKELTCVMKFGSSVGSABRMKEV
 ANLIFSPEDERPVYVLSAMGKTINKLKAGEKAVTGVTVGSLELSFKEHLRTA
 HBLGVETIVIEKHLEGLHOLKGISMKKELTLRTDPLVSFGCKSTRLESAIYNKIG
 HKARQYDAFEIGFTTDDFTNADILKATPVASRTLVGDSKENAVPVATGKGMK
 SCATITLGGSDLTATTIGALGLRBIQWKKVDVLTDPNIYFGAQSVPYLTDE
 AAEIAYFGAQLHPLSRPADGDIPIRVKRSYNPFLPFTVITRSRDSKAVLTIVL
 KRNVTMLDIASRLMGYFLAKVFTTFEDLGSIVDVATSEVSLTLDPALMRE
 LIORNELNIEVERLEKIAVKLIQRSIIISLIGNVKSLLIEKVFQYVRSNVMVQ
 MISOASRKNISLIVNDEBAQCRAHSAFFENDP"

ORIGIN
 Query Match 34.2%; Score 678.4; DB 8; Length 1923;
 Best Local Similarity 69.1%; Pred. No. 6.4e-154;
 Matches 976; Conservative 0; Mismatches 422; Indels 15; Gaps 3;
 QY 260 GGATGGGGGATCAGCTCAGCTGTGATGAAGTTCCGGGGCTCCTCGTGTGGTGGCC 319
 Db 345 GGAAGTGGCAAGATTAACATGTGTAATGAATTTGGCGGTTATCAGTGGATACGA 404
 QY 320 GCGAGATGGCTGAGAGTGGCGGCTCATCTCTGACGTTCCCGAGAGCGCCCGTGTG 379
 Db 405 GAGAGATGMAAGAGTTGCCAATCTTATATTCAGCTTCCAGATGAGAGCCCGCTATT 464
 QY 380 GTTCTCTGCCATGGGAAAAACCAACAACCTTCTCTGCTGAGAGAGGACGATA 439
 Db 465 GTGCTATAGCATGGGAAAGACATTAATAGCTTTTGAAGCTGAGAGAGAGCTGTT 524
 QY 440 GGGTGGAGATATCATGTTTCTGAATTCGAAGATGGAATATGTCACAAAGCCTCAT 499
 Db 525 ACTTGGGATGATCACTAAGATGAGATTTGAAGATTTGAGCTTTATTAAGAGACCTCAT 584
 QY 500 ATCAAGACGTGATGAACCTTGAGCT--TCCAGATCTGTAATACAGCTTTATGAA 556
 Db 585 TTAAGACTGCTCATGAGCTTGGAGTGAACAACGCTTATGAAAAACCTGAGAGGA 644
 QY 557 CTGGAACAATTAAGAGATGCTATGATGAAGAGCTGAGCGCTAGGACAGAGAC 616
 Db 645 CTACACCACTTCTTAAGGATATCAATGATGAAGAGCTAACCTTACGCAAGAGGAC 704
 QY 617 TACCTGTTCATTTGAGAAATGATGTCACAGATTTTTCGTGTTATTTGAACAAA 676
 Db 705 TACTTGATTTATTTGAGAGATGATGCCAAGAGCTTTTCTGTGATCTCAATAAA 764
 QY 677 ATTGCTGTCAGGACGCGCATATGATGATGATTTGTTTTCATTAACAATGATGAA 736
 Db 765 ATTGCGCAAAAGCCGTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 824
 QY 737 TTTGTAATGCGGATATCTTGAAGCAACCTATCTGCTGTGGGAAAGACTTCATGGG 796
 Db 825 TTTACTAATGCTGATATCTTGAAGCAACATACCTGCGCTCAAAAACATGTTGGGC 884
 QY 797 GACTGATACAGATTCACAGCATACCTGTTCTTCTGCTGCTCTTGGGAGAGGCTGAAA 856
 Db 885 GACTGAGCAAGAAATGACAGTCCCTTTTCTTCTGCTGCTGCTGAGAGAGGATGAGG 944
 QY 857 TCTGTGCTGTAATCTTATGAGCCGAGGTGATGATGATGATGATGATGATGATGATGAT 916

	945	TCTTGTGCATCACTACGCTGGGTAAGGGGTGGCAGATGATTGTGACACTCAACGATTGGG	1004
QY	917	AAAGCTTGGGACCTGAGAGAATTCACAGTTATGGAAAAGATGTGATGGTACTTCTGT	976
Db	1005	AAAGGCGTAGGACTCGGGGAATTCAGGTTTGAAAAAGATGGATGAGATTGTGACTTGT	1064
QY	977	GATCCAAATATCTAACCCACATGCAAGAAGCTGTTCCATCTTAAcATTGAAAGGCCACA	1036
Db	1065	GATCTTAACATAATACCCTGGAGCTCATCTGTTCCGACTTAAcGTTCCGATAGGCACT	1124
QY	1037	GAACCTTGCCTAATTTTGGTGCACAGTTTTCATCCCAATTCATGCATGATGAGACCTCTAGAGA	1096
Db	1125	GGCTTGTCTTACTTTGGTGTCCAGAGTGTTCATTCCTACTGTCTATGAGGCCACAGAAAGAT	1184
QY	1097	GGTGATATTCCAGTTAGGGTTAAGAAITTCATCAACCCTTAAAGCTCCAGGCACCTTATT	1156
Db	1185	GGCGACATTCCTGTCAGAGGTTAAAGACCTGACAAcCCCCACTGCTCCAGAACTGTCACT	1244
QY	1157	ACCAGACAAGAAGACATGATGATTAANGTCTGGTTGTACTAATCTAGCAATAGTGTCAAGTCA	1218
Db	1245	ACCAGATCAAGAGATATAGTAA-----GGCTGTGCTAACCGACATTTGTTCTGAAGGT	1298
QY	1217	AATGTCATATGTTGGACATTTGTGAGCACTTCOGATGCTTGTCAGATGTTTTCTGGCA	1276
Db	1299	AATGTGACCATGTTTGACATATGACAAGCACCCGATGCTTGTCGCAATATGTGTTCTTGCC	1358
QY	1277	AAGGTATCAGGATTTTGCTATATTTGAAGATCTATGATATCTGTGATTTGTGTGACCC	1336
Db	1359	AAGGGT-----TTACCAcATTGMAAGATTTAGATATATCGTTGATGTGTGGCACA	1412
QY	1337	AGTGAAGTTAGTGTCTTGTGTCTCACTTGATCCATCAAAGATCTGAGTAAAGGACTGATA	1396
Db	1413	AGTGAAGTTAGTATATCATTTGACATTTGGATCCAGCAAGCTCTGGGGTGAAGAGTTAAT	1472
QY	1397	CAGCAGGCAAGGAACCTTGACCAATGTAAGTTGAAGCTTGAGAAAATAGCAATGTTGCT	1456
Db	1473	CAGCAGATTAATGAACCTGATTAATTTGTGTGABAAGCTTAABAGAATGTCTGTGTGAAA	1533
QY	1457	CTACTTCAGCAGAGGGCGATAATTTCACTTATCGAAAATGTGAGCAATGCTCTGATA	1518
Db	1533	CTACTTCACAGCAGATCAATCATCTCTCTCTCATAGAAATGTTCAGAAATCCTCACACATA	1592
QY	1517	CTAGAAAACCGGAGCGTGTGCTGAGAAAAGTGGGGTTATATGTCAGATGATCTCGCA	1576
Db	1593	TTTGAAGAAGTGTTCCAAGTATTTTGAAGCAATGGAATGAATGTGACAGATGATCTCGAG	1656
QY	1577	GGAGCGTCAAGGTTAACATGTGCTGATAGTCATGATATAGCGATGCAAAAGGCACTCGTA	1633
Db	1653	GGGCGATCTAAGGTAACATCTCATTTGATGATGATATGAMAAGGAGAGCAATGTGTG	1712
QY	1637	GAAGCCCTTCATCAGGCGTCTTTTAAACACAT	1669
Db	1713	AGGGCTCTTCACATCCGCGCTTCTTTTAAAGCCAT	1745
<hr/>			
RESULT 15	AX180354	1658 bp	DNA linear PAT 06-AUG-2001
LOCUS	AX180354	Sequence 15 from Patent WO0146393.	
DEFINITION	AX180354		
ACCESSION	AX180354.1	GI:15132296	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
	Triticum aestivum (bread wheat)		
	Triticum aestivum		
REFERENCE			
AUTHORS	Falco, S. C., Famodu, O. O. and Thorpe, C. J.		
TITLE	Aspartate kinase		
JOURNAL	Patent: WO 0146393-A 15 28-JUN-2001;		
	E.I. DU PONT DE NEMOURS AND COMPANY (US)		
FEATURES	Location/Qualifiers		

[illegible]

QY 1382 AGTAGGGAAGTGTATGACGAGGAGAGTGAATTGACATGATGATTGAAGAGCTTGAGAA 1441
DB 952 AGTCGTGAATTGATTCAGCAG-----GAGCTTGATCATGTGTTGAAGAGCTTGAAAAG 1005
QY 1442 ATAGCAATTGTTGTTCTACTTTCAGCAGAGGGGATTAATTTCATTATCGAAATGTGGAG 1501
DB 1006 ATGCGGTTGTTCAATCTCTACAGCAGATCATCATTTCCCTGATAGGGAAATGTGCAG 1065
QY 1502 CATGCTCTGTATGATAGAAAAGAGGGAGCGTGTGAGGAAAAGTGGGTTAATGTT 1561
DB 1066 AGATCGTCTGTGATTTTGAGAAAGGGGTTCAAATGTTTACGAGAAATGTGTTAATGTT 1125
QY 1562 CAGATGATCTGCAAGAGAGCGTCAAGGTTAACATGTGCTGATAGTCATGATAGCGAT 1621
DB 1126 CAGATGATTTGCAAGAGGCGGTCAAGGTGAACATCTCCTTGTTGATGACAGCGAG 1185
QY 1622 GCAAGGCACTGCTAAGAGCCCTTCATCAGGCGTCTTTGAAGAGATGTCCTATCAAA 1681
DB 1186 GCGAAGCAAGTGCCTGCAAGCCCTCCACTCGGCATTTTGAGAACGGTTTCTTGTCAGAA 1245
QY 1682 GTCGAAGCGAGAGACT 1698
DB 1246 GTAGAGAGAGCGAGCT 1262

Search completed: March 23, 2004, 09:36:41
Job time : 7890 secs

CC production procedures. Nucleic acid encoding aspartate kinase is useful
CC for producing a transgenic plant, by transforming a plant cell with a
CC chimeric gene comprising nucleic acid encoding aspartate kinase operably
CC linked to a regulatory sequence and regenerating a plant from the
CC transformed plant cell. The present sequence is cholic.p002.k6 (FIS) cDNA
CC clone encoding Zea mays aspartate kinase

Sequence 1985 BP; 506 A; 419 C; 535 G; 523 T; 0 U; 2 Other;

Query Match	Score 1983;	DB 5;	Length 1985;
99.9%			

Matches 1985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GCACGAGACAGACGAGAGACTCAAGAAATGGCAATCCAGCTGGATCCGCTGCGCGGC	60
Db	1	GACCCAGACAGACGAGAGAACTCAAGAAATGGCAATCCAGCTGGATCCGCTGCGCGGC	60
Qy	61	CCGCGCGCTCGTTCCGTCGATCACTCCGCGCAGCTCTGGAATATGTTGAGAGACTGCGGTG	120
Db	61	CCGCGCGCTCGTTCCGTCGATCACTCCGCGCAGCTCTGGAATATGTTGAGAGACTGCGGTG	120
Qy	121	CTTCGGTACCCGAAACCGGGCTCGCGGTGCAAGAGGGTGTGCATGGTGGTGCGCACATC	180
Db	121	CTTCGGTACCCGAAACCGGGCTCGCGGTGCAAGAGGGTGTGCATGGTGGTGCGCACATC	180
Qy	181	CACACAGCGTCGCGCCCAAGCAAGCGAGCGCGGTCTTGGGGCGCTGTCTT	240
Db	181	CACACAGCGTCGCGCCCAAGCAAGCGAGCGCGGTCTTGGGGCGCTGTCTT	240
Qy	241	CGAGAGGCTCGGGATGAGAGGATTGGGGGATCAGCTCAGCGTGGTATGAATCGCGGGG	300
Db	241	CGAGAGGCTCGGGATGAGAGGATTGGGGGATCAGCTCAGCGTGGTATGAAGTTCCGGGGG	300
Qy	301	GTCTCTCGGTGTGTGTGGCCGCGAGGATGCTGAGGTGGCGGCGCTCATCTTGACGTTCCC	360
Db	301	GTCTCTCGGTGTGTGTGGCCGCGAGGATGCTGAGGTGGCGGCGCTCATCTTGACGTTCCC	360
Qy	361	CGAGAGCGCCCGGTGTGTGTCTCTGCGCATGGGGAAAAACCAACAACCTTCTCT	420
Db	361	CGAGAGCGCCCGGTGTGTGTCTCTGCGCATGGGGAAAAACCAACAACCTTCTCT	420
Qy	421	TGCTGAGAGAAAGCAGTAGGGTGTGGAGTTATCTAAGTTCTGAAATCGAAGATGGAA	480
Db	421	TGCTGAGAGAAAGCAGTAGGGTGTGGAGTTATCTAAGTTCTGAAATCGAAGATGGAA	480
Qy	481	TATGTCAAAAGCCTTACATATCAAAACGCTGATGAACCTTGACCTTCAAGNATCTGTAA	540
Db	481	TATGTCAAAAGCCTTACATATCAAAACGCTGATGAACCTTGACCTTCAAGNATCTGTAA	540
Qy	541	TAAACCTTTATGAACTGGAGCAACTTTGAAAGGATATGCGTATGATGAAAGATGAC	600
Db	541	TAAACCTTTATGAACTGGAGCAACTTTGAAAGGATATGCGTATGATGAAAGATGAC	600
Qy	601	GCCTAGAGCAGTACTACCTTGTTCATTTGGAGAAATGCAATGCAATGCCAGATTTTTC	660
Db	601	GCCTAGAGCAGTACTACCTTGTTCATTTGGAGAAATGCAATGCAATGCCAGATTTTTC	660
Qy	661	TGCTTTTGAACAAAAATTCGTTGTAAGAGCAACGCAAGTATGAGCAATTTGATTTG	720
Db	661	TGCTTTTGAACAAAAATTCGTTGTAAGAGCAACGCAAGTATGAGCAATTTGATTTG	720
Qy	721	CATTACAACCTGATGAATTTGTAATGCGATATCTTGAAGCAACTATCTGCTGTTC	780
Db	721	CATTACAACCTGATGAATTTGTAATGCGATATCTTGAAGCAACTATCTGCTGTTC	780
Qy	781	GAGAGACCTTACAGGGGACTGGAATCAAGATTCAGCAATACCTGTGTACTGTGGTCT	840
Db	781	GAGAGACCTTACAGGGGACTGGAATCAAGATTCAGCAATACCTGTGTACTGTGGTCT	840
Qy	841	TGGGAAGGCTGGAATCTGTGTCTGTAACTTATAGCCGAGAGTGTATGTAACTTGGAC	900
Db	841	TGGGAAGGCTGGAATCTGTGTCTGTAACTTATAGCCGAGAGTGTATGTAACTTGGAC	900

QY	90	TGCTAACAACCATGGTAAAGCTTGGAGCTGAGAGAAATTCACAGTATGGAAGAGCTTGA	960
Db	901	TGCTACACACCATTGGTAAAGCCTTGGAGCTGAGAGAAATTCACAGTATGGAAGAGCTTGA	960
QY	961	TGGTGTACTTACTTGTGTATCCAAATATCTAACCCATCATGCAAGACTGTTCCTACTTAAAC	1020
Db	961	TGGTGTACTTACTTGTGTATCCAAATATCTAACCCATCATGCAAGACTGTTCCTACTTAAAC	1020
QY	1021	ATTGTGAAGAGGCCACAGAACTTGGCTTATTTGGTCTCAGGTCTTGTGCATCCCAATCGAT	1080
Db	1021	ATTGTGAAGAGGCCACAGAACTTGGCTTATTTGGTCTCAGGTCTTGTGCATCCCAATCGAT	1080
QY	1081	GAGACTCTGCTAGAGAAGGTGATATTCGAGTTAAGGGTTAAGAAATTCATACACCTTAAAGC	1140
Db	1081	GAGACTCTGCTAGAGAAGGTGATATTCGAGTTAAGGGTTAAGAAATTCATACACCTTAAAGC	1140
QY	1141	TCCAGGACCCCTTATATACAGACAAAGAGACATGAGATAAAGGTCTGGTGTACTAATAG	1200
Db	1141	TCCAGGACCCCTTATATACAGACAAAGAGACATGAGATAAAGGTCTGGTGTACTAATAG	1200
QY	1201	CATAGTGTCTCAAGTCAAAATGTCACTATGTTGGACATTTGGAGCATCCGGATCTTGCTCA	1260
Db	1201	CATAGTGTCTCAAGTCAAAATGTCACTATGTTGGACATTTGGAGCATCCGGATCTTGCTCA	1260
QY	1261	GATGTGTTTCTCGCAAGGGTATCAGGATATTTGCTATATGAAGATCTATGATATCTGT	1320
Db	1261	GATGTGTTTCTCGCAAGGGTATCAGGATATTTGCTATATGAAGATCTATGATATCTGT	1320
QY	1321	GGATGTGTGCTTACCAAGTGAAGTTAGTGTTCGTGTCACTTGATCCATCAAAGATCTG	1380
Db	1321	GGATGTGTGCTTACCAAGTGAAGTTAGTGTTCGTGTCACTTGATCCATCAAAGATCTG	1380
QY	1381	GAGTGAAGGAACGTATACAGCAGGCAAGGTGAATCTTGACATGTAGTGAAGAGCTTGAGAA	1440
Db	1381	GAGTGAAGGAACGTATACAGCAGGCAAGGTGAATCTTGACATGTAGTGAAGAGCTTGAGAA	1440
QY	1441	AATAGCAATGTGTCTTACTTACGACAGAGGGCGGATATTTCACTTATGCGAAATGTGGA	1500
Db	1441	AATAGCAATGTGTCTTACTTACGACAGAGGGCGGATATTTCACTTATGCGAAATGTGGA	1500
QY	1501	GCATCGTCTCTGATACTAGAAAAAGACGGGACGTGTCTAGAGAAAAAGTGGGGTTAATGT	1560
Db	1501	GCATCGTCTCTGATACTAGAAAAAGACGGGACGTGTCTAGAGAAAAAGTGGGGTTAATGT	1560
QY	1561	TCGATGATCTCTGCAAGAGCGGTCAAAAGGTTAACATGTCCGTATATGTCATGATAGCGGA	1620
Db	1561	TCGATGATCTCTGCAAGAGCGGTCAAAAGGTTAACATGTCCGTATATGTCATGATAGCGGA	1620
QY	1621	TGCAAAAGGACATCGTAGAAGCCCTTCATCAGGCGTCTTTGAAAGACGATGTCTTATCA	1680
Db	1621	TGCAAAAGGACATCGTAGAAGCCCTTCATCAGGCGTCTTTGAAAGACGATGTCTTATCA	1680
QY	1681	AGTCGAAGGGGAGAACTTACTCTGTGGGCTGATCAAGTAGAGCTTCTGGGTCCAGGGGT	1740
Db	1681	AGTCGAAGGGGAGAACTTACTCTGTGGGCTGATCAAGTAGAGCTTCTGGGTCCAGGGGT	1740
QY	1741	GTTATCTGTTATAGATTCCTCACTGGCTCCATGACGGGATGCGGATGTGATCATTTGATC	1800
Db	1741	GTTATCTGTTATAGATTCCTCACTGGCTCCATGACGGGATGCGGATGTGATCATTTGATC	1800
QY	1801	ATGTGTTTGGCTTGAACAAGTATGTCTTCAGGTTCTCAGCCATATGACTCGAACAATCTGT	1860
Db	1801	ATGTGTTTGGCTTGAACAAGTATGTCTTCAGGTTCTCAGCCATATGACTCGAACAATCTGT	1860
QY	1861	TTCGTGTTTAAAGCTGTTTGCAGACACAGTAGACCTGCGAGCACCGATGTCAACAGAT	1920
Db	1861	TTCGTGTTTAAAGCTGTTTGCAGACACAGTAGACCTGCGAGCACCGATGTCAACAGAT	1920
QY	1921	GGCAACCCCTGTATATATTCCAACTGTCTTAATCAATATATATATATTAACATTATCA	1980
Db	1921	GGCAACCCCTGTATATATTCCAACTGTCTTAATCAATATATATATATTAACATTATCA	1980
QY	1981	TATCT 1985	

Db 1254 TTTTCTGCAAGGATTTGTATAT-----TTGAAGATCTATGATATCTGTGATG 1307
 QY 1327 TGTGTACCACTGAAAGTTAGTGTCTGTGTCATCTTATCCATCAAAAGATCTGAGTAG 1386
 Db 1308 TGTGTCTACAGGAGGTTAGTGTCTGTGTCATCTTATCCATCAAAAGATCTGAGTAG 1367
 QY 1387 GGAAGTATACAGCAGCAGCAAGTGAATTGACCATGATGATTAAGAGCTTGAAGAAATAGC 1446
 Db 1368 GGAAGTATACAGCAG-----GAAGTATGACATGATGATTAAGAGCTTGAAGAAATAGC 1421
 QY 1447 AATTGTCTCTACTTACAGCAGCAGCAATTTTCACTTATCGGAAATGTGAGCATC 1506
 Db 1422 AATTGTCTCTACTTACAGCAGCAGCAATTTTCACTTATCGGAAATGTGAGCATC 1481
 QY 1507 GTCTCTGATCTAGAAAAGACGCGACGTGTGCTGAGAAAAGTGGGTTAATGTTTCAGAT 1566
 Db 1482 GTCTCTCACTAGAAAAGACGCGACGTGTGCTGAGAAAAGTGGGTTAATGTTTCAGAT 1541
 QY 1567 GATCTGCAAGAGCCTCAAGGTTAACAATGCTGTGATGATCCATGATAGCATGCAAA 1626
 Db 1542 GATCTGCAAGAGCCTCAAGGTTAACAATGCTGTGATGATCCATGATAGCATGCAAA 1601
 QY 1627 GGCATCTGATAGAACCCCTCATACAGCGCTTTCTTGAAGACGATGCTATCACAAGTCGA 1686
 Db 1602 GGCATCTGATAGAACCCCTCATACAGCGCTTTCTTGAAGACGATGCTATCACAAGTCGA 1661
 QY 1687 AGCGAGAACCTTACTCTGTGGGCTGATCAACGTAAGCTTGTGGTCCAGCGCTGTTATC 1746
 Db 1662 AGCGAGAACCTTACTCTGTGGGCTGATCAACGTAAGCTTGTGGTCCAGCGCTGTTATC 1721
 QY 1747 TGTATAGATCCCACTGCGCTCCATGAAACGCGATGAGGATGATCATGATGATTT 1806
 Db 1722 TGTATAGATCCCACTGCGCTCCATGAAACGCGATGAGGATGATCATGATGATTT 1781
 QY 1807 TGTGTAACAAGATGCTTCCAGGTTCTCAGCCATGACTGCAAACTGTGTTCTGT 1866
 Db 1782 TGTGTAACAAGATGCTTCCAGGTTCTCAGCCATGACTGCAAACTGTGTTCTGT 1841
 QY 1867 TTTAGAACGTTTGGACACACAGATGAGCTGCGACACCGATTTGTCACAAAGATGGCAAG 1926
 Db 1842 TTTAGAACGTTTGGACACACAGATGAGCTGCGACACCGATTTGTCACAAAGATGGCAAG 1901
 QY 1927 CCTGTGATATTAATTCGAAGTCTCTAATCAATATATATATTAATTAATTC 1978
 Db 1902 CCTGTGATATTAATTCGAAGTCTCTAATCAATATATATATTAATTAATTC 1953
 RESULT 3
 ABZ13939
 ID ABZ13939 standard; DNA; 1710 BP.
 AC ABZ13939;
 XX
 DT 21-JUN-2003 (first entry)
 XX
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1744.
 XX
 KW Arabidopsis thaliana; plant; gene; stress; transgenic; de.
 XX
 OS Arabidopsis thaliana.
 XX
 PN MO200216655-A2.
 PD 28-FEB-2002.
 PF 24-AUG-2001; 2001WO-US026685.
 PR 24-AUG-2000; 2000US-0227866P.
 PR 26-JAN-2001; 2001US-0264647P.
 PR 22-JUN-2001; 2001US-0300111P.
 PA (SCRT) SCRIPPS RES INST
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Harper JF, Krepe J, Wang X, Zhu T;
 XX WPI; 2002-304127/34.
 DR
 XX
 PT Identifying a stress condition to which a plant cell has been exposed and
 PS producing plants with increased tolerance to these abiotic stresses.
 XX
 PS Claim 144; SEQ ID NO 1744; 577bp + Sequence listing; English.
 CC
 CC The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office
 XX
 SO Sequence 1710 BP; 478 A; 323 C; 426 G; 483 T; 0 U; 0 Other;
 Query Match 35.6%; Score 707.6; DB 6; Length 1710;
 Best Local Similarity 70.5%; Pred. No. 1.5e-198;
 Matches 1008; Conservative 0; Mismatches 401; Indels 21; Gaps 4;
 QY 254 ATGAGGAGATTGGGAGATCAGCTCAGCGTGTGATGAAGTTCGGGGGTTCTGCTGTG 313
 Db 235 ACGAGGTAGATGAGAAAGGATACAGCTGCTGATGAAGTGTGGATCTTGTGGGGG 294
 QY 314 TGGGCGGAGAGATGCTGAGATGGCCGCTCATCTCTGACGTTCCCGCAGAGCGCC 373
 Db 295 TCAAGTGAAGAAAGAGAGAGAGTGTGCTGATTTGATTTTGTCTTCCGAGAAAGTCC 354
 QY 374 GTGCTGTTCTCTGCGCATGGGAAACCAACCAACCACTTCTCTGCTGAGAGAG 433
 Db 355 GTCATTTCTCTGCTGATGGGAAACCAACCAATCTTCTGCTGCGGAGAGAG 414
 QY 434 GCAAGTGGTGTGAGATTCATGTTTCTGAAATCGAAGTGGATATGTCATTAAGC 493
 Db 415 GCGGTTAATGTGTGTTCTTAATGATCTGAGTTGAGAGTGTGACATTAATTAAGAA 474
 QY 494 CTACATATCAAGACGATGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 550
 Db 475 TTGCATATCAGAGCGGTGAAGAGCTCAACATGATCCCTGTATATTGACTATTTG 534
 QY 551 TATGAACTGAGCAATCTTGAAGGATATGCTATGATGAAGAGCTGACGCTAGAGCC 610
 Db 535 GAGAACTGAGCAATCTTGAAGGATATGCTATGATGAAGAGCTGACGCTAGAGCC 594
 QY 611 AGTGAATCTGTTGTTGTTGAGAAATGATGCTGCAAGATTTTCTGCTATTTG 670
 Db 595 AAGATTAATAGTCTCTTTGAGAGATGTTTGTCAAGATTTTGTGCTATTTCT 654
 QY 671 AACAATTCGTGTCAAGGACGCGAGTATGACGATTTGATATGTTGTTCAATTAAC 730
 Db 655 AATACAAATCGGTGTCAAGGACGCGAGTATGACGATTTGATATGTTGTTCAATTAAC 714
 QY 731 GATGAATTTGTAATGCGGATATCTTGAAGCAATCTCTGCTGTGGAAGAGCTT 790
 Db 715 GATGATTTCAAAATGAGGATATCTTGAAGCAATCTTCTGCTGTGGAAGAGCTT 774
 QY 791 CATGGGACTGATACAGATTCACGATATCTTGTATCTGTGCTTCTTGGAAAGGC 850
 Db 775 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 834
 QY 851 TGGAAATCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 910
 Db 835 TGGAAATCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894

QY 911 ATTGTAAGCCTTGGGACTGAGAGAAATTCAGGATGAGAAATGTTGATGCTGACT 970
DB 895 ATTGTAAGCCTTGGGACTGAGAGAAATTCAGGATGAGAAATGTTGATGCTGACT 954
QY 971 ACTTGATCCAAATATCTACCAATCCAAAGACTGTTCCATCTTAACATTTGAAG 1030
DB 955 ACATGTGACCTACTATATATAGGAGCTACACCAATACCATCTGACATTTGAGAA 1014
QY 1031 GCCACAGAACTTCTATTTTGTCTGCTGAGTTTGCATCCAAATGATGAGACTGCT 1090
DB 1015 GCAGCGAGCTAGCTTATTTTGTGACAGGCTTGCACCAATGATGATGAGCAGCA 1074
QY 1091 AAGAAAGTGAATTTCCAGTTAGGTTAAGATTTATCAACCTTAAGCTCCAGGACC 1150
DB 1075 AAGAGAGGTGAGATTCCTGTTAGGTTAAATTTTAAACCTTAAGCTCCAGGACT 1134
QY 1151 CTTATACCAAGCAAAAGACATGATTAAGGTCTGTTGTTACTACTAGCATAGCTCTC 1210
DB 1135 ATCATCTCTAAACAAAGACATGACCAAGS-----TATTTAAGAGCATTTGCTG 1188
QY 1211 AAGTCAATGCACTATGTTGACATTTGTGACACTGAGATGCTTGTCTGATGCTTT 1270
DB 1189 AAACGTATGATGACATGCTGATATGACAGACCCCAATGCTTGTCAAGTTGCTTT 1248
QY 1271 CTGGCAAGGTATCAGTATTTGCTATATGTAAGATCTATGATATCTGTGATTTGTT 1330
DB 1249 CTTGCAAGGTATTTCTGATAT-----TTAGAGCTTGTGCAATTTCCGTATGTTGTT 1302
QY 1331 GCTACCAAGTAAATGATTTCTGTGCTGATCTGATCCATCAAAATCTGAGTAGGAA 1390
DB 1303 GCACACTAGGAAGTCAATCTCTGACACTGATCTTCAAACTTTGAGAGAGAA 1362
QY 1391 CCGATCAGCAGGCAAGTAACTGACCACTGATGTAAGAGCTTGAAGAAATGCAAT 1450
DB 1363 CTGATCAACAG-----GAGCTTGATCTGATGTTGAAGACTGAGAAATTTGCACTT 1416
QY 1451 GTTCCTTACTTACAGAGAGGCGATTAATTTCACTTATCGAAATGTGAGCAATGCT 1510
DB 1417 GTGAATCTTAAAGAGAGCAATCATCTCTTAATGGGAATGTTCAATCTCTCC 1476
QY 1511 CTGATCTGAAGAAAGCGGAGCTGTGCTGAGAGAAAGTGGGTTAAGTTCAGATATC 1570
DB 1477 CTGATTTAGAGAGGCGCTTATGTTCTTAATACCAAGGTAAATGTCAGATGATA 1536
QY 1571 TCGCAGAGAGCGTCAAGGTAACTATGCTGATAGTCCATGATGAGATGCAAGCA 1630
DB 1537 TCGCAGAGAGCATCCAGGTAAATCTTCTTATAGTAAAGCAAGCTGAAGCTGAAGT 1596
QY 1631 CTGCTAGAGCCCTTATCAGGCGTTCTTTGAAGCATGCTCTATCA 1680
DB 1597 TGTGTTCAGGCTTTCACAAATCTTCTTTCGAGAGCGGTGACCTCTCGA 1646

RESULT 4

AAC37959 standard; DNA; 1933 BP.

AAC37959;

17-OCT-2000 (first entry)

Arabidopsis thaliana DNA fragment SEQ ID NO: 19281.

XX Hybridisation assay; genetic mapping; gene expression control;
XX Protein identification; signal transduction pathway; metabolic pathway;
XX Promoter; termination sequence; ss.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130448P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132565P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134321P.
PR 14-MAY-1999; 99US-0134370P.
PR 14-MAY-1999; 99US-0134376P.
PR 18-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139892P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144842P.
PR 20-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147326P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148365P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149712P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151865P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 12-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.

PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159283P.
PR 13-OCT-1999; 99US-0159284P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160982P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 35.0%; Score 695.6; DB 3; Length 1933;
Best Local Similarity 70.4%; Pred. No. 578-195;
Matches 1007; Conservative 0; Mismatches 401; Indels 22; Gaps 5;

QY 254 ATGAGAGATTGGGGGATCAGCTCAGCGTGTGATGAAAGTTCGGGGGTCCTCGTGTG 313
DB 280 ACGGAGTATGATGAAAGGATCAGCGTGTGATGAAAGTTCGGTGTGATGAAAGG 339
QY 314 TGGCCCGGAGAGATGCGTGAAGTGGCCGCTATCTGTACGTTCCCGAGAGCGCCC 373
DB 340 TCAGCTGAGAAAGAAAGAAAGTGTGATTTGACTTTCCGAAAGAAAGTCCC 399
QY 374 GTTCGCTTCTCTCTGACCATGGGAAACACACACACTTCTCCTTGTGGAGAG 433
DB 400 GTCAATTCTCTCTCTGCTATGGGAAACACACACTTCTTGTGGAGAGAG 459
QY 434 GCAGTAGGATGATGATTCATCTTCTGAAATCGAAGATGAAATGTCACAAAGC 493
DB 460 GCGGTTAGTGTGGTGTCTTCAATGCACTGATGATGAGATGAGCAATTAAAGAA 519
QY 494 CTACATATCAGACGGTGGATGAAGTGGACT---TCCAGNATCTGTATACAGCTT 550
DB 520 TTGCATATCAGACGGTGAAGAAAGCTCAACATTGATCCCTGTATTTGACTATTT 579
QY 551 TATGAATGAGCAACTTATGAAGATGCTATGATGAAGAGCTGACGCTTAGACC 610
DB 580 GAGGAATGAGCAACTCTGAAGGCAATGCCATGAT-AAAGATGACACTTGAACC 638
QY 611 AGTGAATGCTGTTTCTTATTTGAGAAATGATGCTCCAGGATTTTGTGCTATTTG 670
DB 639 AGAGATTACTTGTCTTTTGGAGAGTGTGTCTCAAGAGATTTTGTGCTATTTG 698
QY 671 AACCAATTCGTGTACAGCAGGAGATGACCAATTTATGTTGCTTCAATTAACCT 730
DB 699 AATCAATTCGTGTACAGCAGGAGATGACCAATTTATGTTGCTTCAATTAACAG 758

PR	23-JUN-1999;	99US-0139893P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141877P.	PR	13-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	16-SEP-1999;	99US-0154039P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	08-JUL-1999;	99US-0142820P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142877P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	13-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144684P.	PR	18-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160981P.
PR	26-JUL-1999;	99US-0145276P.	PR	22-OCT-1999;	99US-0160989P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145919P.	PR	25-OCT-1999;	99US-0161406P.
PR	28-JUL-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161359P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146389P.	PR	26-OCT-1999;	99US-0161361P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161920P.
PR	04-AUG-1999;	99US-0147302P.	PR	28-OCT-1999;	99US-0161922P.
PR	05-AUG-1999;	99US-0147260P.	PR	28-OCT-1999;	99US-0161933P.
PR	06-AUG-1999;	99US-0147303P.	PR	29-OCT-1999;	99US-0162142P.
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147959P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			

QY	CTGGACAACTATTGAAAGATCGCTATGATGAAAGCTGACGCTGAGCACTGAC	616
Db	55 CTGGACAACTATTGAAAGATCGCTATGATGAAAGCTGACGCTGAGCACTGAC	616
QY	CTACACCACTCTCTAAAGGCATATCATGATGAAAGGTAACTTACGCAAGGAC	775
Db	716 CTACACCACTCTCTAAAGGCATATCATGATGAAAGGTAACTTACGCAAGGAC	775
QY	TACCTGTTTCACTTTGGAAATGCAATGCAAGATTTTTCCTGTTATTGAAACA	676
Db	617 TACCTGTTTCACTTTGGAAATGCAATGCAAGATTTTTCCTGTTATTGAAACA	676
QY	TACTTGTTTCACTTTGGAGAGCATGCTCCAAAGCTTTTCTCTCGATCTCATATA	835
Db	776 TACTTGTTTCACTTTGGAGAGCATGCTCCAAAGCTTTTCTCTCGATCTCATATA	835
QY	ATTCCGTGAAGGCACGGCAGATGATGACGATTTGATATGGCTTCACTTACATGATGA	736
Db	677 ATTCCGTGAAGGCACGGCAGATGATGACGATTTGATATGGCTTCACTTACATGATGA	736
QY	ATTGGCCAAAGACCGGTCAATATGATGCAATTTGATCGAATTTTACCCACAGACAC	895
Db	836 ATTGGCCAAAGACCGGTCAATATGATGCAATTTGATCGAATTTTACCCACAGACAC	895
QY	TTTGGTAATCCGATATCTTAAAGAACCACTATCCGTGTTGCGAAGACCTTCATGGG	796
Db	737 TTTGGTAATCCGATATCTTAAAGAACCACTATCCGTGTTGCGAAGACCTTCATGGG	796
QY	TTTACTAATGCTGATATATCTTGAAGCAACATACCTCCGCTCCAAATAACCTTGTGGC	955
Db	896 TTTACTAATGCTGATATATCTTGAAGCAACATACCTCCGCTCCAAATAACCTTGTGGC	955
QY	GACTGGATACAGATATCCACGCAATACCTGTGTGTTCTGTGGGTTCTTTGGAAAGGGCTGGAAA	856
Db	797 GACTGGATACAGATATCCACGCAATACCTGTGTGTTCTGTGGGTTCTTTGGAAAGGGCTGGAAA	856
QY	GACTGGAGCAAGAAATATGCAATCCCTGTGTGTTTCTGTGCACTCGGAAAAGGAGTAGAG	1015
Db	956 GACTGGAGCAAGAAATATGCAATCCCTGTGTGTTTCTGTGCACTCGGAAAAGGAGTAGAG	1015
QY	TCGTGTCGTAACTACTTTAGGCCGAGTGTGATCACTTGACTCTCAACCACTAGT	916
Db	857 TCGTGTCGTAACTACTTTAGGCCGAGTGTGATCACTTGACTCTCAACCACTAGT	916
QY	TCCTTGTCATATCACTACCTGGTGGATGGGATGCGATATTTGACAGCTTCAACGATTTGG	1075
Db	1016 TCCTTGTCATATCACTACCTGGTGGATGGGATGCGATATTTGACAGCTTCAACGATTTGG	1075
QY	AAAGCCTTGGAGCTGAGAAATTCAGGATGGAAGATGTTGATGTTACTTACTTCTGT	976
Db	917 AAAGCCTTGGAGCTGAGAAATTCAGGATGGAAGATGTTGATGTTACTTACTTCTGT	976
QY	AAAGCCCTTGAAGCTCGGGAAATTCAGGTTTGGAAATGTCGATGAGATTTTGTGACTGT	1135
Db	1076 AAAGCCCTTGAAGCTCGGGAAATTCAGGTTTGGAAATGTCGATGAGATTTTGTGACTGT	1135
QY	GATCCAAATATCAATCCCAATGCAAGCAAGCTGTGCAACTTAACTTGAAGAAGCCACA	1036
Db	977 GATCCAAATATCAATCCCAATGCAAGCAAGCTGTGCAACTTAACTTGAAGAAGCCACA	1036
QY	GATCTTAAATATATCTCGAGCTCAATCTGTCCGATCTTAACTTGAAGCTTGAAGCACT	1135
Db	1136 GATCTTAAATATATCTCGAGCTCAATCTGTCCGATCTTAACTTGAAGCTTGAAGCACT	1135
QY	GAACTTGTATTATTTGTGTGTGAGTTTGTGATCCAAATCGATCGATGAGCACTGTAAGA	1096
Db	1037 GAACTTGTATTATTTGTGTGTGAGTTTGTGATCCAAATCGATCGATGAGCACTGTAAGA	1096
QY	GACCTTGCTTACTTTGTGTGTGAGTTTGTGATCGATCGATGAGCACTGTAAGAAT	1255
Db	1196 GACCTTGCTTACTTTGTGTGTGAGTTTGTGATCGATCGATGAGCACTGTAAGAAT	1255
QY	GGTGAATATTCAGTTAAGGTTAAAGATTCAATCAACCTTAAAGCTCCAGGCACTTATT	1156
Db	1097 GGTGAATATTCAGTTAAGGTTAAAGATTCAATCAACCTTAAAGCTCCAGGCACTTATT	1156
QY	GGGGAATATTCGTGAGGGTTAAAGATCGTCAACCCCTGTCTCAGAAATCTGTATC	1315
Db	1256 GGGGAATATTCGTGAGGGTTAAAGATCGTCAACCCCTGTCTCAGAAATCTGTATC	1315
QY	ACCAGCAAAAGAGACATGATTAAGGTTGTGTTTAACTAATACTGATGCTCAAGTCA	1216
Db	1157 ACCAGCAAAAGAGACATGATTAAGGTTGTGTTTAACTAATACTGATGCTCAAGTCA	1216
QY	ACCAGATCAAGAGCATGATGA-----GGCTGTCTAACAGCATTTGTTGAAAGT	1369
Db	1316 ACCAGATCAAGAGCATGATGA-----GGCTGTCTAACAGCATTTGTTGAAAGT	1369
QY	AATGTCAATATGTTGACATTTGTGACCTCGATGCTTGTGTCATATGTTTCTGGCA	1276
Db	1217 AATGTCAATATGTTGACATTTGTGACCTCGATGCTTGTGTCATATGTTTCTGGCA	1276
QY	AATGTACCATGTTGACATGACAGCACCCGTATGTTGGCCAAATATGTTTCTTCTCC	1429
Db	1370 AATGTACCATGTTGACATGACAGCACCCGTATGTTGGCCAAATATGTTTCTTCTCC	1429
QY	AGGGATCAAGGATTTTGATATTTGAAATCTAATGATATCTGTGATGTTGTGCTAC	1336
Db	1277 AGGGATCAAGGATTTTGATATTTGAAATCTAATGATATCTGTGATGTTGTGCTAC	1336
QY	AAAGGT-----TTTACCATTTTGAAGATTTAGGTATATCGGTTATGTTGTGCAACA	1483
Db	1430 AAAGGT-----TTTACCATTTTGAAGATTTAGGTATATCGGTTATGTTGTGCAACA	1483
QY	AGTGAAGTTAGTGTCTGTGTGTGCACTTGTATCCATCAAAAGATCTGAGTAGGAGACTGATA	1396
Db	1337 AGTGAAGTTAGTGTCTGTGTGTGCACTTGTATCCATCAAAAGATCTGAGTAGGAGACTGATA	1396
QY	AGTGAAGTTAGTATATCATATGACATTTGATCCAAAGACCTCGGGGTGAGAGATTATTT	1543
Db	1484 AGTGAAGTTAGTATATCATATGACATTTGATCCAAAGACCTCGGGGTGAGAGATTATTT	1543
QY	CACAGAGCAAGGAACTTGACCATATGATTTGAAGAGCTTGAAGAAATATACATTTTGGT	1456
Db	1397 CACAGAGCAAGGAACTTGACCATATGATTTGAAGAGCTTGAAGAAATATACATTTTGGT	1456
QY	CACAGAGTAAATGAATCTGATATATTTGGTGGAAAGCTGAAGAAAGATCTGTGTGTGAAA	1603
Db	1544 CACAGAGTAAATGAATCTGATATATTTGGTGGAAAGCTGAAGAAAGATCTGTGTGTGAAA	1603
QY	CTACTTCCAGCAGAGGGCGATATTTTCACTTATCCGAATGTGAGCAATCGTCTGTGATA	1516
Db	1457 CTACTTCCAGCAGAGGGCGATATTTTCACTTATCCGAATGTGAGCAATCGTCTGTGATA	1516
QY	CTACTTCCAAAGCAATCAATCATCTCTCTCATTAAGAAATGTTCAAGAAATCCCACTGATA	1663
Db	1604 CTACTTCCAAAGCAATCAATCATCTCTCTCATTAAGAAATGTTCAAGAAATCCCACTGATA	1663
QY	CTTGAAAAGACGGACGTGTGTGAGAAAAGTGGGTTATATGTTCAATGATCTTCGCA	1576
Db	1517 CTTGAAAAGACGGACGTGTGTGAGAAAAGTGGGTTATATGTTCAATGATCTTCGCA	1576
QY	TTTGAAGAAAGGTGTTCCAAAGTATTTTGAAGCAATGATGATGATGATCTTCGAG	1733
Db	1664 TTTGAAGAAAGGTGTTCCAAAGTATTTTGAAGCAATGATGATGATGATCTTCGAG	1733
QY	GGAGGCTCAAAAGTTTAACTGTGCTGATAGCTATATGCAATCCAAAGCACTGCTGA	1636
Db	1577 GGAGGCTCAAAAGTTTAACTGTGCTGATAGCTATATGCAATCCAAAGCACTGCTGA	1636
QY	GGTGATCTTAAGTTAAACATCTCATTTGATATGTAATGATGAAGCAAGCAATGTGTG	1783
Db	1724 GGTGATCTTAAGTTAAACATCTCATTTGATATGTAATGATGAAGCAAGCAATGTGTG	1783
QY	GAAAGCCCTTCAATCAGCGCTTT	

```

Db      1784 AGGGCTCTCCACTCGCTTTTGAGACCGAT 1816
RESULT 6
AAD08626
AAD08626 standard; cDNA, 1658 BP.
AC
AD08626;
AA08626;
04-SEP-2001 (first entry)
DE      Wheat aspartate kinase encoding wri.pk0046.b11 (FIS) cDNA clone.
XX
XX      Wheat; aspartate kinase; cell transformation; transgenic plant;
KW      wri.pk0046.b11 clone; ss.
OS      Trilicium aestivum.
XX
XX      Key      Location/Qualifiers
FT      CDS      1..1320
FT              /tag= a
FT              /product= "Trilicium aestivum wri.pk0046.b11 (FIS)
FT              aspartate kinase"
FT              /note= "CDS does not include start codon"
FT              /partial
XX
XX      WO200146393-A2.
XX
XX      28-JUN-2001.
XX
XX      19-DEC-2000; 2000MO-US034396.
XX
XX      21-DEC-1999; 99US-0172944P.
XX
XX      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX      Falco SC, Farneda OO, Thorpe CJ;
PI
DR      WPI; 2001-418057/44.
DR      P-PSDB; AA04358.
XX
XX      Novel nucleic acid molecules encoding aspartate kinase useful in cell
PT      transformation and transgenic plant production procedures.
XX
XX      Claim 8; Page 59; 64pp; English.
XX
XX      The present invention relates to nucleic acid molecules encoding
CC      aspartate kinase useful in cell transformation and transgenic plant
CC      production procedures. Nucleic acid encoding aspartate kinase is useful
CC      for producing a transgenic plant, by transforming a plant cell with a
CC      chimeric gene comprising nucleic acid encoding aspartate kinase operably
CC      linked to a regulatory sequence and regenerating a plant from the
CC      transformed plant cell. The present sequence is cholic.pk002.k6 (FIS) cDNA
CC      clone encoding Trilicium aestivum aspartate kinase
XX
XX
XX      Sequence 1658 BP; 457 A; 331 C; 417 G; 453 T; 0 U; 0 Other;
SQ
Query Match      34.0%; Score 674.6; DB 5; Length 1658;
Best Local Similarity 72.8%; Pred. No. 8.8e-189;
Matches 930; Conservative 0; Mismatches 326; Indels 21; Gaps
OY      425 GGAGAGAAAGCGAGTGGGTGAGATTATCCATGTTTCTGAATCGAAGAGTGAATATG 484
Db      4 GGAGACAGGCGCGTGAAGTGGGCGGCCCCCAAGCGCTCGGAATCTAGAGCTCGCCGTC 63
OY      485 GTCAAAAGCTTCATATCAAGACGCTGATGAACCTTGAGACTTCCAAAGATCTGTAAATCA 544
Db      64 ATCAAGAGACTCCATCTCGAGACCATGATGAGAGCTTGCGCTAGATGCTCCATTGTTTCA 123
OY      545 AGCGTTT---ATGAAGCTGAGCAACTATTTGAAGATATCGCTATGATGAAAGCTGACG 601
Db      124 GGTTTTGTGAGCAATTGAGCAACTGCTCAAGGATCTTCTATGATGAAAGCTGACT 183

```


QY 602 CCTAGGACGAGTACTTCTTTCATTTGGAGATGATGTCACCGAGATTTTCT 661
 Db 184 CTAGAGACAGGATTAACCTTGTTCCTTTGGTGAATGATGTCACAGAAATTTTCT 243
 QY 662 GCTTATTTGAACAAATTCGTGTCAAGGACGGGAGTGAAGCATTTGATTTGTTCT 721
 Db 244 GCATTTTGAATTAATCTAGGAGAGAGGACGACGATTAATGCTTTGATCTTGCTTT 303
 QY 722 ATTACACTGATGAATTTGTAATGCGATATCTTGAAGCAACCTATCTGCTGTCG 781
 Db 304 ATAAACCATGACGATTTCAAAATGCCGATATCTCGAAGCAACTTATCCTGCTGTGA 363
 QY 782 AAGAGCTTATGGGAGCTGATACAGATTCAGGAGTACCTGTTTACTGGGTCTT 841
 Db 364 AAGAGCTTATGAGATTTGATGATGACCTGCTATTTCTTAATGATGCTGTTCTT 423
 QY 842 GGGAGGCTGGAATCTGTGCTGTAACTATTAGCCGAGGTGATGACTGACT 901
 Db 424 GGGAGGATGGAATCTGTGGGTCAACAAGTTAGAGAGGGGCGGAGTGACTGACC 483
 QY 902 GCTACAACTATGTTGTAAGCCCTTGGAGCTGAGAGAAATTCAGTATGGAAGATTTGAT 961
 Db 484 GCTACAACTATGGAAGCTTGGGTTTAAAGAAATTCAGGTTTGAAGAGATGTAGAC 543
 QY 962 GGTGACTTACTTGTGATCCAAATATCTACCCAGATGGAAGACTGTTCCATCTTACA 1021
 Db 544 GGTGTTGACGTTGATCCAAATATTTATGCAAGCGGTACAGTACCTACTTGACT 603
 QY 1022 TTTGAAGAGCCACAGAACTTGTATTTTGTGTCTGAGTTTGGATCCCAATCGATG 1081
 Db 604 TTTGATGAGGAGCTGAACTTGTATTTTGTGACAGGTTTGGATCCCAATCGATG 663
 QY 1082 AGACCTGTAGAGAAAGTATATTCAGTTAGGTTAAGATTCAACACCTTAAGCT 1141
 Db 664 CAGCCAGCAAGGAGAGTGTATCCAGTTCAGTGAAGAACTCATTAACCGTCAAGCA 723
 QY 1142 CCAGGACCCCTTATTAACAACAAGACATGATTAANGTCTGTGTCTACTATGAC 1201
 Db 724 CCGGCACTGTGATCTATAAACAGAGATATGCCCAAG-----CATATTAACAGC 777
 QY 1202 ATAGTCTAAGCAATGCACTATGTGACATTTGAGACTCGATCGATCTGTGTAG 1261
 Db 778 ATTTGCTGAATCAAAATTTACATGCTGATTAATGAGACAGAGATGCTGGACAG 837
 QY 1262 TATGTTTTCTGGAAGGATTCAGGATTTGCTATATTGAAGATCTATGATATCTGTG 1321
 Db 838 TATGCTTTCTAGCAAGGCTTCTCAATAT-----TTGAAGATTTGGGTATCTGTGTT 891
 QY 1322 GATTTGTTGCTACCGTAAGTATGTTTCTGTGATCACTGATCCATCAAAAGTCTGG 1381
 Db 892 GATTTGTTGCTACCGTAAGTATGATGATGATGATGATGATGATGATGATGATGATG 951
 QY 1382 AGTAGGAACTGATACAGAGGCAAGTGAACCTGACCATGATTTGAAGACTTGAGAA 1441
 Db 952 AGTGTGAATGATTCAGAG-----GAGCTTGAATGATGATTTGAAGACTTGAGAA 1005
 QY 1442 ATAGCAATTTGTTGTTACTTTCAGAGAGGCGATTAATTTCTATGCGAAATGTGAG 1501
 Db 1006 ATAGCGTTGTTATCTCTACAGCAATCAATCAATTTCTCTATGAGGAAATGTGAG 1065
 QY 1502 CAATGCTCTGATCTAGAAAAGACGGGAGCTGTGCTGAGAAAAGTGGGTTATGTT 1561
 Db 1066 AGATCGTCTGATCTTAGGAAAGCGGCTTCAATGTTCTACGAGAAATGTGATATGTT 1125
 QY 1562 CAATGATCTGGAAGAGAGCTCAAGTTAATGATGCTGATGATGATGATGATGATGATG 1621
 Db 1126 CAGATGATTTGGAAGAGGCGCTCAAGGATGATCTCTGTTGATGATGATGATGATGATG 1185
 QY 1622 GCAAGGCACTGCTAGAGCCCTTCAATCAAGGCTTCTTGAAGAGATGCTCATCAACA 1681
 Db 1186 GCGAAGGAGTGGCGTCAAGCCCTTCACTGCGGATTTTGAAGAGGTTTCTTGTCAAGA 1245

QY 1682 GTCGAGCGGAGACT 1698
 Db 1246 GTAGGAGAGCGGACT 1262
 RESULT 7
 AAC41888
 ID AAC41888 standard; DNA; 1395 BP.
 XX AAC41888;
 XX 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 33506.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway; metabolic pathway;
 KW promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 28-APR-1999; 99US-0130891P.
 PR 30-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132464P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0135621P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137502P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138647P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.

QY 380 GTTCTCTGCGATGGGGAAAAACCAACCACTTCTCTGCTGGAGAGAGCAGTA 439
 Db 541 GTGCTATACGAAATGGGAAAGACATATATAGCTTTTGAAGCTGGAGAGAGCTGT 600
 QY 440 GGGTGTGAGTATTCATGTTTCTGAATCGAAGTGAATATGTCAAAAGCCTTACT 499
 Db 601 ACTTCGGGTGCATTAAGTGAAGATTTGAAGATTTAGCTTTTAAAGAACTCCAT 660
 QY 500 ATCAGAGCGTGAATGAATCTTGACT--TCCAGNATCTGTATATACAGCTTTTGA 556
 Db 661 TTAAAGACTGCTCATAGCTTTGAGTGAACAAACGGTTATTTGAATAACCTTGAAGA 720
 QY 557 CTGAGCAACTATTTGMAAGGTATGCTATGATGAAGAGCTGACGCTTAGACCTGAC 616
 Db 721 CTACACCAAGCTTCTCAAGGCAATATCATGATGAAGAGCTTAACCTTACGCAAGGGAC 780
 QY 617 TACCTTGTTCATTTTGGAGATGATGCTCCACCGAATTTTCTGTTATTTGAACAA 676
 Db 781 TACTTGGTTTCAATTGGAGAGTGAATGCCAAGAGCTTTTCTTGGTATCTCATTA 840
 QY 677 ATTGCTCAAGGCAAGGAGATGACGATTTGATATATGTTTCAATTAACACTGATA 736
 Db 841 ATTGCCCAAAAGCGGCTCAATATGATGATTTGATGATCGGATTTATCACCAAGAGAC 900
 QY 737 TTTGGTAATCGGATATCTTGAAGCAACCTATCTGCTGTGGAGAGACTTCATGGG 796
 Db 901 TTTACTAATGCTGATATCTTGAAGCAACATACCTGCGCTCAAAAACACTTGTGGC 960
 QY 797 GACTGATACAGATCCAGCATCCTGTTGTAAGTCTTGGTCTTGGAAAGGCTGAAA 856
 Db 961 GACTGAGCAAGGAAATGCAAGTCCCTGTTGTAAGTCTTGGTCTTGGAAAGGCTGAGG 1020
 QY 857 TCTGCTGCTAATCTTACCTTTAGGCGGAGGTGTAGTGAATCTGCTCAACCATTTGT 916
 Db 1021 TCTTGTGCTATCACTACGCTGGGTAGGGGTGCGAGTGTGACAGCTACACCATTTGG 1080
 QY 917 AAAGCCTGGAGCTAGAGAAATTCAGGTATGAAAGATGTTGATGTTACTTACTTGT 976
 Db 1081 AAAGCCTAGAGCTGCGGAAATTCAGGTATGAAAGATGTTGATGTTGACTTGT 1140
 QY 977 GATCCAAATATCTACCCATGCAAGACTGTTCCATCTTAACTTTGAAGAGCCACA 1036
 Db 1141 GATCTTAATATATCTGAGACTCAATCTGTTCCGTATTAAGTTCGATGAGAGCGCT 1200
 QY 1037 GAACCTGCTATTTGGTCTAGGTTTGCATCCACATGATGAGAGCTGCTAGAGAA 1096
 Db 1201 GAGCTTGCTTACTTTGGTCCAGAGTGTGATCCACTGCTTAGAGGCCAGC-ATAGAT 1259
 QY 1097 GGTGATATTCAGTTAGGTTAAGATTCATACAACTTAAAGCTCCAGGCACTTATT 1156
 Db 1260 GGGGACATTCCTGTAGGGGTAAAGACTGTTACAACTCCAGCTCCAGAACTGTCATC 1319
 QY 1157 ACCAGCAAGAGACATGATATANGTCTGTTGATACATAGATATGTTGCTCAAGTCA 1216
 Db 1320 ACCAGATCAAGAGACATGATTAAGCTTTG-----CTAACGACATTTCTGAAAGCT 1373
 QY 1217 AATGCACTATGTTGACATTTGAGCACTGAGCTGTTGCTAGTATGTTTCTGGCA 1276
 Db 1374 AATGCACTATGTTGACATTTGAGCACTGAGCTGTTGCTAGTATGTTTCTGGCC 1433
 QY 1277 AAGGTATCAAGTATTTGCTATATGAAGTCTATATATCTGAGATGTTGTTGCTACC 1336
 Db 1434 AAGGTATTTTACCACTTTGAGATTTAGGATATGCTTTGATGTTGCTGCAACA 1487
 QY 1337 AGTGAAGTATGTTTCTGCTCACTTGAATTCATCAAAAGTCTGAGATGAGAACTGATA 1396
 Db 1488 AGTGAAGTATGTTTCTGCTCACTTGAATTCATCAAAAGTCTGAGATGAGAACTGATA 1545
 QY 1397 CAGCAGCAAGTGAATTTGACATGATGATGAGAGCTTGAAGAAATGCAATTTGTTGCT 1456
 Db 1546 CAGCAGCAAGTGAATTTGATGATGATGATGAGAGCTTGAAGAAATGCTTATTTGTA 1605

QY 1457 CTACTTACGAGAGGGCGATATTTCACTTATTCGMAAATGTGAGCAATCTCTCTGATA 1516
 Db 1606 CTACTTACGAGAGATCAATCACTCTCTCATAGGMAAATGTGCAAAATCTCACTCATTA 1665
 QY 1517 CTGAAAGAGCGGACGCTGCTGAGAGAAAGTGGGTTAATGTTCAAGTATCTCGCA 1576
 Db 1666 TTAGAGAGGCTGTTCCAGATATTTAGAGCAATGAGATGATGCAATGATCTCGCAG 1725
 QY 1577 GGAAGCTCAAGGTTAATGCTGCTGATGCTGATGATGATGATGATGATGATGAT 1636
 Db 1726 GGTGATCTTAAGTTAAACATCTCATGATGATGATGATGATGATGATGATGATGAT 1785
 QY 1637 GAGCCTTCATGAGGCTTCTTGAAGCAT 1669
 Db 1786 AGGCTCTCCATCTCGCTTCTTGGATGAT 1818
 RESULT 8
 AAC32977
 ID AAC32977 standard; DNA; 2164 BP.
 XX AAC32977;
 AC 17-OCT-2000 (first entry)
 DT 17-OCT-2000 (first entry)
 XX 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 1340.
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 1340.
 KM Hybridisation assay; genetic mapping; gene expression control;
 KM promoter identification; signal transduction pathway; metabolic pathway;
 KW promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX EP1033405-A2.
 PD 06-SEP-2000.
 XX 06-SEP-2000.
 XX 25-FEB-2000; 2000EP-00301439.
 PF 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126746P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0136293P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139692P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141827P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.

PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150565P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156566P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.0%; Score 654.6; DB 3; Length 2164;
 Best Local Similarity 68.9%; Pred. No. 8,9e-183;
 Matches 964; Conservative 0; Mismatches 415; Indels 21; Gaps 4;

QY GGGGATGACGCTGAGTGTGATGAGTTCGGGGGGTCTCGGTGTCGGCGCGAG 325
 Db GGGGATGACGCTGAGTGTGATGAGTTCGGGGGGTCTCGGTGTCGGCGCGAG 325
 QY 559 GAGAAAAATTGACTGTGTATGAAATTCGGTGCATCAGTGCATCAGCAAGG 618
 Db 559 GAGAAAAATTGACTGTGTATGAAATTCGGTGCATCAGTGCATCAGCAAGG 618
 QY 326 ATGCTGAGTGGCGGCTCATCTGACGTTCCCGAGAGCGCCCGTGTCTTC 385
 Db 326 ATGCTGAGTGGCGGCTCATCTGACGTTCCCGAGAGCGCCCGTGTCTTC 385
 QY 386 TGTGCTATGGGAAAAACCAACCACTTCTCCTGTGTGAGAGAGGAGTGGT 445
 Db 386 TGTGCTATGGGAAAAACCAACCACTTCTCCTGTGTGAGAGAGGAGTGGT 445
 QY 679 TCAGCAATGGCAAGACCAACATAGCTTTGATGCTGGGGAGAGGCTTTGCTGC 738
 Db 679 TCAGCAATGGCAAGACCAACATAGCTTTGATGCTGGGGAGAGGCTTTGCTGC 738
 QY 446 GGAGTTATCATGTTTCTGAAATCGAAGATGGAATATGATCAAAAGCTTACATATCAAG 505
 Db 446 GGAGTTATCATGTTTCTGAAATCGAAGATGGAATATGATCAAAAGCTTACATATCAAG 505
 QY 739 GGTGTACCAACGTGACACTATCGAAGGTTGACCTATATTAAGAACTCCATATAGG 798
 Db 739 GGTGTACCAACGTGACACTATCGAAGGTTGACCTATATTAAGAACTCCATATAGG 798
 QY 506 ACCGTGATGAACTTGACTTCCAGATTCGTATAC--AAGCTTATGAGCTGAG 562
 Db 506 ACCGTGATGAACTTGACTTCCAGATTCGTATAC--AAGCTTATGAGCTGAG 562
 QY 799 ACTGCTCATGAGCTTGGATGGAAACAGCACTATTGCAAACTTGAAGAGCTGGA 858
 Db 799 ACTGCTCATGAGCTTGGATGGAAACAGCACTATTGCAAACTTGAAGAGCTGGA 858
 QY 563 CAACTATTGAAGATATGCTATGATGAAGAGCTGACGCTAGAGCACTACTACTT 622
 Db 563 CAACTATTGAAGATATGCTATGATGAAGAGCTGACGCTAGAGCACTACTACTT 622
 QY 859 CAGCTTCTTAAGGGGTGGCATGATGAAGAAATTAATTACGCTGAAGAGCTACTTG 918
 Db 859 CAGCTTCTTAAGGGGTGGCATGATGAAGAAATTAATTACGCTGAAGAGCTACTTG 918
 QY 623 GTTTCATTGGAAATGCATGTCACCGAGATTTTCTGCTTATTGAAACAAATTCGT 682
 Db 623 GTTTCATTGGAAATGCATGTCACCGAGATTTTCTGCTTATTGAAACAAATTCGT 682
 QY 919 GTTTCATTGGAAATGCATGTCACCGAGATTTTCTGCTTATTGAAACAAATTCGT 978
 Db 919 GTTTCATTGGAAATGCATGTCACCGAGATTTTCTGCTTATTGAAACAAATTCGT 978
 QY 683 GTCAAGGCAAGGAGATGACGATTTGATGATGATGATGATGATGATGATGATGAT 742
 Db 683 GTCAAGGCAAGGAGATGACGATTTGATGATGATGATGATGATGATGATGATGAT 742
 QY 979 CACAAAGCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038
 Db 979 CACAAAGCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038
 QY 743 AATGCGATATCTTGAAGCAACCTATCTGCTGTGGAAGAGACTTCATGCGGACTCG 802
 Db 743 AATGCGATATCTTGAAGCAACCTATCTGCTGTGGAAGAGACTTCATGCGGACTCG 802
 QY 1039 AATGCTGATATCTTGAAGCAACCTATCTGCTGTGGAAGAGACTTCATGCGGACTCG 1098
 Db 1039 AATGCTGATATCTTGAAGCAACCTATCTGCTGTGGAAGAGACTTCATGCGGACTCG 1098
 QY 803 AATAGGATCCAGGATCCTGTGTACTGGGCTCCTGGAAGAGGCTGGAATTCGT 862
 Db 803 AATAGGATCCAGGATCCTGTGTACTGGGCTCCTGGAAGAGGCTGGAATTCGT 862
 QY 1099 AGCAGGAAAAATGCACTCCCGCTGTGACAGGCTTCTTGAAGAGAGTGAATCTTGT 1158
 Db 1099 AGCAGGAAAAATGCACTCCCGCTGTGACAGGCTTCTTGAAGAGAGTGAATCTTGT 1158
 QY 863 GCTGTACTACTTATGAGCGAGGTGTGATGATGATGATGATGATGATGATGATGAT 922
 Db 863 GCTGTACTACTTATGAGCGAGGTGTGATGATGATGATGATGATGATGATGATGAT 922
 QY 1159 GCTGTACTACTTATGAGCGAGGTGTGATGATGATGATGATGATGATGATGATGAT 1218
 Db 1159 GCTGTACTACTTATGAGCGAGGTGTGATGATGATGATGATGATGATGATGATGAT 1218
 QY 923 TTGGACTGAGAAATTCAGGTATGAGAAAGTGTGATGATGATGATGATGATGATGAT 982
 Db 923 TTGGACTGAGAAATTCAGGTATGAGAAAGTGTGATGATGATGATGATGATGATGAT 982
 QY 1219 TTAGGTTTGGGAGATCCAGGTTGGAAGATGAGATGAGATGATGATGATGATGAT 1278
 Db 1219 TTAGGTTTGGGAGATCCAGGTTGGAAGATGAGATGAGATGATGATGATGATGAT 1278
 QY 983 AATATCTACCAATGCAAGACTGTTCTCATCTTACATTTGAAGAGCCACAACTT 1042
 Db 983 AATATCTACCAATGCAAGACTGTTCTCATCTTACATTTGAAGAGCCACAACTT 1042
 QY 1279 AACATTTACTGCGAGACTCAACCTGTTCCACCTTGAATGATGAGGCTGCAACTT 1338
 Db 1279 AACATTTACTGCGAGACTCAACCTGTTCCACCTTGAATGATGAGGCTGCAACTT 1338
 QY 1043 GCTTATTTGGTGTCAAGTTTGGATCCCAATGAGAGAGACCTGCTAGAGAGTGT 1102
 Db 1043 GCTTATTTGGTGTCAAGTTTGGATCCCAATGAGAGAGACCTGCTAGAGAGTGT 1102
 QY 1339 GCTTACTTTGGTGTCAAGTTTGGATCCCAATGAGAGAGACCTGCTAGAGAGTGT 1398
 Db 1339 GCTTACTTTGGTGTCAAGTTTGGATCCCAATGAGAGAGACCTGCTAGAGAGTGT 1398
 QY 1103 ATTCAAGTTAGGTTTGAATTCATACAACTTAAGCTCCAGGACCTTATTACAGA 1162
 Db 1103 ATTCAAGTTAGGTTTGAATTCATACAACTTAAGCTCCAGGACCTTATTACAGA 1162
 QY 1399 ATTCTGTAGGTTTGAATTCATACAACTTCCAGGACCTTATTACAGA 1458
 Db 1399 ATTCTGTAGGTTTGAATTCATACAACTTCCAGGACCTTATTACAGA 1458
 QY 1163 CAAAGACATGATATANGTCTGTGTGATTAATTAAGAGTCTCAATCAATGTC 1222
 Db 1163 CAAAGACATGATATANGTCTGTGTGATTAATTAAGAGTCTCAATCAATGTC 1222
 QY 1459 TCAAGACATGAGCA-----GGCTGATTGACAGAGTGTCTTAAAGTAAATGTC 1512
 Db 1459 TCAAGACATGAGCA-----GGCTGATTGACAGAGTGTCTTAAAGTAAATGTC 1512
 QY 1223 ACTATGTTGACATTTGTGAGCACTCGAGTGTGTGATGATGATGATGATGATGATGAT 1282
 Db 1223 ACTATGTTGACATTTGTGAGCACTCGAGTGTGTGATGATGATGATGATGATGATGAT 1282
 QY 1513 ACCATGTTGAGCATCTAGCAACCGGTATGCTCGGCAATATGATGATGATGATGATGATGAT 1572
 Db 1513 ACCATGTTGAGCATCTAGCAACCGGTATGCTCGGCAATATGATGATGATGATGATGATGAT 1572

QY 1283 TCAGTATTTGCTATATGAAAGATCTATATATCTGTGATGATGATGATGATGATGATGAT 1342
 Db 1573 T-----TTTCCATTTGAAAAATGGGCAATCTGTGATGATGATGATGATGATGATGAT 1626
 QY 1343 GTTATGTTTCTGTGCTACTGATTCATCAAGATTTGAGTGGAACTGATACAGCAG 1402
 Db 1627 GTTATGATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1685
 QY 1403 GCAAGTAACTTGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1462
 Db 1686 -----CGAGCTTGTATGAGTGTGAGAGCACTTGAAGAAATGCTTGTAAATCTGCTT 1740
 QY 1463 CAGAGAGGAGGATATTTCACTATGGAATGTGAGCAATGCTCTGATACTAGAA 1522
 Db 1741 CGACAGATCAATATATCTCTCATGGAATGTTCAAGATCATCATATATAGAG 1800
 QY 1523 AAGAGGAGCTGTGCTGAGAGAAAGTGGGTTATGTTCAATGATGATGATGATGATGAT 1582
 Db 1801 AAGGCTTCCAGATCTTCAACCAATGGGATTAATGTCAGATGATGATGATGATGATGAT 1860
 QY 1583 TCAAAGTTAATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1642
 Db 1861 TCTAAGTAAACATCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
 QY 1643 CTTATCAGGCTTCTTGA 1662
 Db 1921 CTCACCTGGGCTTCTTGA 1940

RESULT 9
 AAC48615
 ID AAC48615 standard; DNA; 1692 BP.
 AC AAC48615;
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 58119.
 XX
 XX Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway; metabolic pathway;
 KW promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000BP-00301439.
 XX
 PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0133180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0128845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.
PR 11-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134210P.
PR 18-MAY-1999; 99US-013470B.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138440P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139899P.
PR 22-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 23-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 29-JUN-1999; 99US-0141287P.
PR 30-JUN-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142890P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-014431P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150889P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0156563P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0158294P.
PR 13-OCT-1999; 99US-0158295P.
PR 14-OCT-1999; 99US-0158329P.
PR 14-OCT-1999; 99US-0158330P.
PR 14-OCT-1999; 99US-0158331P.
PR 14-OCT-1999; 99US-0158637P.
PR 14-OCT-1999; 99US-0158638P.
PR 18-OCT-1999; 99US-0158584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.

PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 31.0%; Score 616.2; DB 3; Length 1692;
 Best Local Similarity 66.6%; Pred. No. 1.9e-171;
 Matches 979; Conservative 0; Mismatches 419; Indels 72; Gaps 4;

260 GGATTGGGGGATCAGCTGAGCGTGTGATGAGTGGGGGGTCCCTGGTGTGCGGCGC 319
 229 GGAATGGCAAGAGTTGACATGTGTAAATGAAATTTGGCGTTTCATCAGTGAATACGA 288
 320 GCGAGGATGCTGAGGTGGCGGCTCATCTGACGTTCCCGAGAGCGCCCGTCTGTC 379
 289 GAGAGGATGAAAGATGTGCCATCTTACTCAGCTTTCAGATGAGAGCCGTCATT 348
 380 GTTCTCTGCGCATGGGGGAAACACCAACCACTTCTCTGCTGGAGAGAGGAGCA 439
 349 GTGCTATCAGCAATGGGAAAGACGACTAATGACTTTTGAAGGCTGAGAGAGGCTGTT 408
 440 GGGTGTGAGTTATCCATGTTTCTGAATCGAAGAGTGAATGTGCAAAAGCCATCAT 499
 409 ACTTGGCGGTGCTACGTAAGTGAAGATTTGAAGATTTGAAGGATTTAAAGAACTCCAT 468
 500 A-----TC 502
 469 TTAAGTGTCAAGTCCAAATTTAGCCTTATTTCTCATGGAAATPAAATTTCTTATGTC 528
 503 AAGACGATGATGAATCTGCAATTCGAAGATCTGTAAATCAAG---CCTTTATGAAGT 559
 529 AGAATGCTCATGAGCTTGGAGTGAAGAAACAGGTTATGAAAACACCTGAAAGAGCA 588
 560 GAGCACTATTGAAAGAGTATGCTATGATGAAAAGCTGACGCTGAGAGCACTGACTAC 619
 589 CACCAAGCTTCCAAAGGATATCAATGATGAAAGGCTAACCTTACGCAAGAGGACTAC 648
 620 CTGTTTCATTTGAGAGATGATGATCCACAGATTTTCTGCTTATTTGAACAAAT 679
 649 TTGTTTCATTTGAGAGTGTGATGCCAAGGCTTTCTCTGGTATCTCAATAAAT 708
 680 CGTGTCAAGGCAAGGCAAGTATGACGATTTGATTTGGTTTCATTAACATGATGATTT 739
 709 GAGCACAAGGCGGCTCATATGATGCAATGAGATCGAATTTATACCAAGACGACTTT 768
 740 GATATGCGGATATCTTGAAGCAACCTATCTCTGCTGGAAGAGACTTCAATGGGAGC 799
 769 ACTATGCTGATATCTTGAAGCAACATACCTGCGCTCCCAAAAACACTGTGTGCGAGC 828
 800 TGATACAGATCCAGCGATACCTGTGTTTACTGTGTTTCTTGGGAAAGGCTGGAATCT 859
 829 TGGAGCAAGGAAATGCGAGTCCCTGTGTTTACTGCGTACCTCGAAAGGAGTGAAGTCT 888
 860 GATGCTTAATCTTATTTAGCGGAGGTGTGATGATGATGATGATGATGATGATGATGAT 919
 889 TGTGTATCATCACTGCTGAGTGGAGGAGTGTGATGATGATGATGATGATGATGATGAT 948
 920 GCTTGGAGCTGAGAGAAATTCAGGTATGAGAGATGATGATGATGATGATGATGATGAT 979
 949 GCGTAGGAGCTGCGGGAATTCAGGTATGAGAGATGATGATGATGATGATGATGATGAT 1008
 980 CCAATATCTATCCCAATGAGCAAGAGCTGTTCCATTAATCAATTTGAAGAGGCAAGAA 1039
 1009 CCTAATATATCTCGAGGCTCATCTGTTCCGACTTAAGCTTCATGATGAGGAGAGCTGAG 1068

QY 1040 CTTCCTATTTTGTGCTCAGGTTTTCATCCAAATGATGAGACTGTGAGAGAGT 1099
 DB 1069 CTTCCTATTTTGTGCTCAGGTTTTCATCCAAATGATGAGACTGTGAGAGAGT 1128
 QY 1100 GATTTTCATTTAGGTTTAAAGATTCATCAACCTTAAAGCTTCCAGGCACTTTTAC 1159
 DB 1129 GACTTCTCTGAGGTTTAAAGCTGTGACCAACCTGCTCCAGGAACTGTCTATCAC 1188
 QY 1160 AGACAAAGACATGATTAAGTCTGTTGTTACTAATGATGATGATGATGATGATGAT 1219
 DB 1189 AGATCAAGACATGATTA-----GCTGTGCTTACAGGATTTCTGAAAGTAT 1242
 QY 1220 GTCACTATTTGACATTTGAGACATCGGATGCTTGTGCTGATGATGATGATGATGAT 1279
 DB 1243 GTGACCATTTGACATGACAGACACCGTATGCTTGGCCAAATGATGATGATGATGAT 1302
 QY 1280 GTATCAGGATTTGCTATATTTGAAGATCTATGATATCTGATGATGATGATGATGAT 1339
 DB 1303 GTGT-----TACCAATTTGAAGATTTAGATATCGGTTGATGATGATGATGATGAT 1356
 QY 1340 GAAGTTATGTTTCTGTGCTATTTTCCATCAAGATCTGAGTATGAGGAACTGATAC 1399
 DB 1357 GAAGTTATGATATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1416
 QY 1400 CAGCAAGTGAATCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
 DB 1417 CGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1476
 QY 1460 CTTCAGCAGAGGCGATATTTTCTATTCGAAATGTGAGCAATCTGCTGTGATCTA 1519
 DB 1477 CTTCAGCAGAGATCAATCTCTCTCATGAGAAATGTTCAAAATCTCTCACTCATATTA 1536
 QY 1520 GAAAGAGGAGGAGCTGTGCTGAGAGAAAGTGGGTTAATGTCAGATGATCTCGCAAGA 1579
 DB 1537 GAGAGGTTTTCAGATTTTGAAGACATGAGAGTGAAGATGATGATGATGATGATGAT 1596
 QY 1580 GCGTCAAGGTTTAAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1639
 DB 1597 GCATCTAAGGTTAAATCTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1656
 QY 1640 GCCCTCATGAGGCTTCTTGAAGAGAT 1669
 DB 1657 GCTTCCATCTCGGCTTCTTGAAGAGAT 1686

RESULT 10
 AAD08619
 ID AAD08619 standard; cDNA; 565 BP.
 XX
 AC AAD08619;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Zea mays aspartate kinase encoding bmsl.pK0008.e5 cDNA clone.
 XX
 KW Corn; aspartate kinase; cell transformation; transgenic plant;
 KW bmsl.pK0008.e5 clone; ss.
 OS
 XX
 FH Zea mays.
 XX
 Key location/Qualifiers
 CDS 2..295
 FT /*tag= a
 FT /product= "Zea mays bmsl.pK0008.e5 aspartate kinase"
 FT /transl_except= (pos:125..127, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid; CDS
 FT does not include start codon"
 FT /partial
 XX
 PN MO200146393-A2.
 XX
 PD 28-JUN-2001.

PF 19-DEC-2000; 2000WO-US034396.
 XX
 PR 21-DEC-1999; 99US-0172944P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Falco SC, Famodu OO, Thorpe CJ;
 XX
 XX WPI: 2001-418057/44.
 DR P-PSDB; AAE04351.
 XX
 XX Novel nucleic acid molecules encoding aspartate kinase useful in cell
 PT transformation and transgenic plant production procedures.
 XX
 XX Claim 8; Page 47; 64pp; English.
 CC The present invention relates to nucleic acid molecules encoding
 CC aspartate kinase useful in cell transformation and transgenic plant
 CC production procedures. Nucleic acid encoding aspartate kinase is useful
 CC for producing a transgenic plant, by transforming a plant cell with a
 CC chimeric gene comprising nucleic acid encoding aspartate kinase operably
 CC linked to a regulatory sequence and regenerating a plant from the
 CC transformed plant cell. The present sequence is bmsl.pk0008.es cDNA clone
 CC encoding Zea mays aspartate kinase
 XX
 SQ Sequence 565 BP; 158 A; 115 C; 137 G; 154 T; 0 U; 1 Other;

Query Match 28.0%; Score 555; DB 5; Length 565;
 Best Local Similarity 98.9%; Pred. No. 1.4e-153;
 Matches 558; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1418 CATGATGTTGAAGAGCTGGAGAAATAGCAATTTGCTCTTACTTACGAGAGGCGATA 1477
 DB 2 CACGAGGTTGAAGAGCTGGAGAAATAGCAATTTGCTCTTACTTACGAGAGGCGATA 61
 QY 1478 ATTTCACTTATCGGAAATGTTGAGCAATCGTCTCTGATCTAGAAAAGACGGACGTGTG 1537
 DB 62 ATTTCACTTATCGGAAATGTTGAGCAATCGTCTCTGATCTAGAAAAGACGGACGTGTG 121
 QY 1538 CTGAGAAAAGTGGGGTTATGTTCAATGATCTCGCAAGAGGCTCAAGGTTAATGATG 1597
 DB 122 CTGAGAAAAGTGGGGTTATGTTCAATGATCTCGCAAGAGGCTCAAGGTTAATGATG 181
 QY 1598 TCGCTGATGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1657
 DB 182 TCGCTGATGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 241
 QY 1658 TTTGAAGAGATGCTCTATCACAAGTCGAAGCGGAGAACTTACTGTGGGCTGATCAACG 1717
 DB 242 TTTGAAGAGATGCTCTATCACAAGTCGAAGCGGAGAACTTACTGTGGGCTGATCAACG 301
 QY 1718 TAGGCTTGGCTGGGTCGAGGCGGTATCTGTTAGATTCGCACTGCGCTCCATGAACG 1777
 DB 302 TAGGCTTGGCTGGGTCGAGGCGGTATCTGTTAGATTCGCACTGCGCTCCATGAACG 361
 QY 1778 GCATGGGCAATTTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1837
 DB 362 GCATGGGCAATTTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421
 QY 1838 AACCAATGATCTGCAAAACCTGTTCTGTTTGAAGCTGTTTGAAGCAGCAGAGAGCTG 1897
 DB 422 AACCAATGATCTGCAAAACCTGTTCTGTTTGAAGCTGTTTGAAGCAGCAGAGAGCTG 481
 QY 1898 CGAGCACCGATGTCACAAGATGCAAGCTGTCGATATTAATTCATGCTCTTAATCA 1957
 DB 482 CGAGCACCGATGTCACAAGATGCAAGCTGTCGATATTAATTCATGCTCTTAATCA 541
 QY 1958 AT 1981
 DB 542 AT 565

RESULT 11

AAD08620
 ID AAD08620 standard; cDNA; 513 BP.
 XX
 AC AAD08620;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 XX Zea mays aspartate kinase encoding choic.pk002.k6 cDNA clone.
 DE Corn; aspartate kinase; cell transformation; transgenic plant;
 KW choic.pk002.k6 clone; ss.
 XX
 XX Zea mays.
 OS
 XX
 XX Key Location/Qualifiers
 FH 1..459
 FT CDS
 FT /*tag= a
 FT /product= "Zea mays choic.pk002.k6 aspartate kinase"
 FT /note= "CDS does not include start codon"
 FT /partial
 PN WO200146393-A2.
 XX
 XX 28-JUN-2001.
 PD
 PF 19-DEC-2000; 2000WO-US034396.
 XX
 XX 21-DEC-1999; 99US-0172944P.
 PR
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Falco SC, Famodu OO, Thorpe CJ;
 PI WPI: 2001-418057/44.
 XX P-PSDB; AAE04352.
 DR
 XX Novel nucleic acid molecules encoding aspartate kinase useful in cell
 PT transformation and transgenic plant production procedures.
 XX
 XX Claim 8; Page 48; 64pp; English.
 PS The present invention relates to nucleic acid molecules encoding
 CC aspartate kinase useful in cell transformation and transgenic plant
 CC production procedures. Nucleic acid encoding aspartate kinase is useful
 CC for producing a transgenic plant, by transforming a plant cell with a
 CC chimeric gene comprising nucleic acid encoding aspartate kinase operably
 CC linked to a regulatory sequence and regenerating a plant from the
 CC transformed plant cell. The present sequence is choic.pk002.k6 cDNA clone
 CC encoding Zea mays aspartate kinase
 XX
 SQ Sequence 513 BP; 102 A; 140 C; 170 G; 100 T; 0 U; 1 Other;

Query Match 22.6%; Score 448.6; DB 5; Length 513;
 Best Local Similarity 96.6%; Pred. No. 4.6e-122;
 Matches 479; Conservative 0; Mismatches 15; Indels 2; Gaps 2;

QY 8 ACAGAGCAGAGAGATCTGAAATGCAATCCAGTGGATGCGCTGCGCGCCCGCCG 67
 DB 1 ACAGAGCAGAGAGATCTGAAATGCAATCCAGTGGATGCGCTGCGCGCCCGCCG 60
 QY 68 CTCGTTCCGTGATACCTCCGCGGAGCTGCGATGTTGAGAGACTGCGGCTTCG 127
 DB 61 CTCGTTCCGTGATACCTCCGCGGAGCTGCGATGTTGAGAGACTGCGGCTTCG 120
 QY 128 ACCCGAACCGGCTCGCGGATGCAAGAGGTTGCAATGATGATGATGATGATGATGATG 187
 DB 121 ACCCGAACCGGCTCGCGGATGCAAGAGGTTGCAATGATGATGATGATGATGATGATG 180
 QY 188 CGTGGGCGAAGCAAGCGGAGCGGCGGAGCGCTCTTGGGCGGCTGTTCTCGAGG 247
 DB 181 CGTGGGCGAAGCAAGCGGAGCGGCGGAGCGCTCTTGGGCGGCTGTTCTCGAGG 240
 QY 248 CTCGGAGTGAAGGATGCGGAGTCACTGACGCTGATGATGAAGTTGCGGCGGCTCTCG 307


```

Db      241 CTGGGATGAGAGGATTGGGGGATCAGCTGAGCTGTGATGAGTTTGGGGGGTCTCTG 300
Qy      308 GTGTGTGTCGCGCGAGAGATGAGTGTGAGTGGCCGAGCTCACTCTGACGTTCCCGAGAG 367
Db      301 GTGTGTGTCGCGCGAGAGATGAGTGTGAGTGGCCGAGCTCACTCTGACGTTCCCGAGAG 360
Qy      368 GCGCCCGTCGTCTCTCTCTGCAATGAGGAAACCAACCAACCTTCTCTTGTCT-GG 426
Db      361 GCGCCCGTCGTCTCTCTCTGCAATGAGGAAACCAACCAACCTTCTCTTGTCTGAG 420
Qy      427 AGAGAGGACAGTGTGAGTGTGATTCATGTTTCTGAAAT-CGAGAGTGTGAAATATG 485
Db      421 AGAAAGGCAATTAAGTGTGAGTGTATCATGTTTCTGAAATCGAAGAAATGAAATATG 480
Qy      486 TCAAAAGCCTACATAT 501
Db      481 TCAAAAGCCTAAAT 496

RESULT 12
AAB08624
ID      AAB08624 standard; cDNA; 847 BP.
AC      AAB08624;
XX
XX      04-SEP-2001 (first entry)
DT
XX      Oryza sativa aspartate kinase encoding rd1f.pk005.f20 (FIS) cDNA clone.
DE
XX      Rice; aspartate kinase; cell transformation; transgenic plant;
KM      rd1f.pk005.f20 clone; ss.
XX
XX      Oryza sativa.
OS
XX      Location/Qualifiers
FH      3. 845
FT      /tag= a
FT      /product= "Oryza sativa rd1f.pk005.f20 (FIS) aspartate
FT      kinase"
FT      /note= "CDS does not include stop codon"
FT      /partial
XX
XX      W0200146393-A2.
XX
XX      28-JUN-2001.
PD
XX
XX      19-DEC-2000; 2000WO-US034396.
PF
XX
XX      21-DEC-1999; 99US-0172944P.
PR
XX
XX      (DUP0 ) DU PONT DE NEMOURS & CO E. I.
PA
XX
XX      Falco SC, Famodu OO, Thorpe CJ;
PI
XX
XX      WPI; 2001-418057/44.
DR      P-PSDB; AAE04356.
XX
XX      Novel nucleic acid molecules encoding aspartate kinase useful in cell
XX      transformation and transgenic plant production procedures.
XX
XX      Claim 8; Page 55-56; 64pp; English.
XX
XX      The present invention relates to nucleic acid molecules encoding
XX      aspartate kinase useful in cell transformation and transgenic plant
XX      production procedures. Nucleic acid encoding aspartate kinase is useful
XX      for producing a transgenic plant, by transforming a plant cell with a
XX      chimeric gene comprising nucleic acid encoding aspartate kinase operably
XX      linked to a regulatory sequence and regenerating a plant from the
XX      transformed plant cell. The present sequence is rd1f.pk005.f20 (FIS)
XX      cDNA clone encoding Oryza sativa aspartate kinase
XX
XX      Sequence 847 BP; 180 A; 196 C; 270 G; 201 T; 0 U; 0 Other;
SQ

```

```

Query Match      18.7%; Score 371.2; DB 5; Length 847;
Best Local Similarity 73.8%; Pred. No. 5,7e-99;
Matches 485; Conservative 0; Mismatches 169; Indels 3; Gaps 1;

Qy      274 GCTCAGCGTGTGATGAGTGTGAGGAGGTCCTGTGAGTGTGTGCGCGAGAGATGCTGA 333
Db      182 GTTACCGTGTCCATTAAGTGTGCGGAGTGTGTGTGCGCGAGAGATGCTGA 241
Qy      334 GTGTGCGGCGCTATCTGACGTTCCCGAGAGAGCGCCGTCGTCTTCTCTGCAAT 393
Db      242 GTGTGCGGAGTCTATCTGACGTTCCCGAGAGAGATGCTGTTGTCTCTCCGCAAT 301
Qy      394 GGGGAAACCAACCAACCAACCTTCTCTGAGAGAGAGAGTGTGAGGATGAGTAT 453
Db      302 GGGGAAACCAACCAACCAACCTTCTCTGAGAGAGAGAGTGTGAGGATGAGTAT 361
Qy      454 CCATGTTTCTGAATGAAAGTGAATATGTCTCAAAAGCCTACATATCAAGACGTGA 513
Db      362 GAAAGCGCTCTGAATTCGCCAGCTCCAGATTAACAAGAGCTCATGTAGACTATTA 421
Qy      514 TGAAGTGTGATTCAGAGATTCGTATTAACAGCCTT---ATGAAGTGAAGCACTAT 570
Db      422 TGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 481
Qy      571 GAAAGTATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
Db      482 TAAAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541
Qy      631 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 690
Db      542 CGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
Qy      691 ACGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 750
Db      602 TCGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
Qy      751 TATCTTGAAGCAACCTATCTCTGTCGAGAGAGACTCTATGGGACTGATACAGA 810
Db      662 CATTTTGAAGCAACCTATCTCTGTCGAGAGAGACTCTATGGGACTGATACAGA 721
Qy      811 TCCAGCATACCTGTTGTTACTGAGGTTCTCTGAGAGAGGCTGGAATCTGAGCTGTAAC 870
Db      722 CCGGATATCTATAGTACTGTTCTCTGAGAGAGGATGAGATCATGCTGCTCAC 781
Qy      871 TACTTTAGCGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 927
Db      782 CACTTTAGCGAGAGTGTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 838

RESULT 13
ABX84348
ID      ABX84348 standard; cDNA; 282 BP.
AC      ABX84348;
XX
XX      24-APR-2003 (first entry)
DT
XX
XX      Corn ear-derived polynucleotide (cpd) #2808.
DE
XX
XX      Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022;
XX      structural gene; functional gene; regulatory gene;
XX      corn ear-specific profile; gene transcription; gene expression;
XX      hybrid plant; desirable trait expression; plant breeding program;
XX      inheritance; desired characteristic; growth; development;
XX      disease resistance; environmental adaptability; quality; yield;
XX      multigene trait; plant; gene; ss.
XX
XX      Zea mays.
OS
XX
XX      US6476212-B1.
XX
XX      05-NOV-2002.
PD

```

	XX	PF	14-MAY-1999;	99US-00313294.	
	XX	PR	26-MAY-1998;	98US-0086722D.	
	XX	PA	(INCY-) INCYTE GENOMICS INC.		
	XX	PI	Lalajudi RV, Ito LY, Sherman BK;		
	XX	PT	WPI; 2003-208840/20.		
	XX	PS			
	XX	XX	Example; SEQ ID NO 2808; 390bp; English.		
PT	CC	CC	The present invention relates to the isolation of corn ear-derived		
PT	CC	CC	polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022		
PT	CC	CC	and SATMON023. Some of the cdps uniquely identify structural, functional,		
PT	CC	CC	and regulatory genes of corn ear. The polynucleotide sequences are		
PT	CC	CC	useful for detecting cdps in a sample, for producing a corn ear-specific		
PT	CC	CC	profile of gene transcription, for detecting altered gene expression in		
PT	CC	CC	inbred or hybrid plants, and for screening several molecules for specific		
PT	CC	CC	binding to the polynucleotide. The cdps are useful to identify, isolate,		
PT	CC	CC	or extend identical or related corn-ear nucleic acid sequences from DNA		
PT	CC	CC	libraries, and in nucleic acid amplification or hybridisation techniques		
PT	CC	CC	to follow the expression of desirable traits through plant breeding		
PT	CC	CC	programs. Preferably, the cdps are used to identify, evaluate, alter, or		
PT	CC	CC	follow the inheritance of desired characteristics associated with growth		
PT	CC	CC	and development, disease resistance, environmental adaptability, quality,		
PT	CC	CC	and yield of corn. The cdps are also useful as molecular markers for		
PT	CC	CC	studying inheritance and multigene traits in a plant breeding program.		
PT	CC	CC	The cdps are useful for producing purified corn-ear polypeptides by		
PT	CC	CC	recombinant techniques. They are also useful in diagnostic assays to		
PT	CC	CC	detect or confirm conditions or diseases associated with abnormal levels		
PT	CC	CC	of cdp expression. ABX81541-ABX89140 represent corn ear-derived		
PT	CC	CC	polynucleotides (cdps) of the invention. Note: The sequence data for this		
PT	CC	CC	patent did not form part of the printed specification, but was obtained		
PT	CC	CC	in electronic format directly from the USPRO web site at		
PT	CC	CC	seqdata.uspro.gov/psidsidentry.html		
SO	XX	XX	Sequence 282 BP; 87 A; 50 C; 64 G; 81 T; 0 U; 0 Other;		
Query Match			12.3%; Score 244.2; DB 7; Length 282;		
Best Local Similarity			97.1%; Pred. No. 1.4e-61;		
Matches 271; Conservative			0; Mismatches 3; Indels 5; Gaps 2;		
DG	YY	YY	449 GTATTCACATTCTTGCAGAAATCGAAGAAGAGGAATTGGCAAAAAGCCTCATATCAAGACG 508		
DG	DB	DB	1 GTTATTCACATTCTTGCAGAAATCGAAGAAGAGGAATTGTGTCAAAAGCCTCATATCAAGACG 60		
QY	YY	YY	509 GTGGATGAACCTTGACCTTCCAAGNACTGTAAACAAG---CCTTTAAGACTGAGACA 564		
QY	DB	DB	61 GTGGATGAACCTTGACCTTCCAAG-ATCTGTAAATCAAAGACATGTAGATNACTGAGACA 119		
QY	YY	YY	565 ACTATTGAAGAAGATACGCTATGATTAATAAAGACTGACGCGTAGACCAAGTACACTGCTGT 624		
DB	DB	DB	120 ACATTTGAAGAAGATACGCTATGATTAATAAAGACTGACGCGCTTAGACCAAGTACACTGCTGT 179		
QY	YY	YY	625 TTCAATTGAGAAATGCATGTCCACCAGGATTTTTTTGTGCTTAATTGACAAAATTCGTGT 684		
DB	DB	DB	180 TTCAATTGAGAAATGCATGTCCACCAGGATTTTTTTGTGCTTAATTGACAAAATTCGTGT 239		
QY	YY	YY	685 CAAGCACGCGACGATGAGACGACTTTGATTAATGAGTTTCAT 723		
DB	DB	DB	240 CAAGCACGCGACGATGAGACGACTTTGATTAATGAGTTTCAT 278		
RESULT 14					
ABX843393					
ID					
ABX843393 standard; cDNA; 270 BP.					

AC	XX	ABX84393;
AD	XX	
AE	XX	24-APR-2003 (first entry)
AF	XX	
AG	XX	Corn ear-derived polynucleotide (cpd) #7853.
AH	XX	
AI	XX	Corn ear-derived polynucleotide; cdp; CDNA library; SATMON022; SATMON023;
AJ	XX	structural gene; functional gene; regulatory gene;
AK	KW	corn ear-specific profile; gene transcription; gene expression;
AL	KW	hybrid plant; desirable trait expression; plant breeding program;
AM	KM	inheritance; desired characteristic; growth; development;
AN	KM	disease resistance; environmental adaptability; quality; yield;
AO	XX	multigene trait; plant; gene; ss.
AP	OS	Zea mays.
AQ	PN	US6476212-BI.
AR	XX	
AS	PD	05-NOV-2002.
AT	XX	
AV	PF	14-MAY-1999; 99US-0033294.
AW	PR	26-MAY-1998; 98US-0086722P.
AX	PA	(INCY-) INCTE GENOMICS INC.
AY	XX	Lalgudi RV, Ito LY, Sherman BK;
AZ	DR	WPI; 2003-208840/20.
BA	PT	Novel purified corn-ear derived polynucleotide useful as hybridization
BB	PT	probe for detecting polynucleotide in sample, and for identifying,
BC	PT	evaluating, and altering desired characteristics associated with growth,
BD	PT	development.
BE	PS	Example; SEQ ID NO 2853; 390bp; English.
BF	XX	
BG	CC	The present invention relates to the isolation of corn ear-derived
BH	CC	polynucleotides (cdps) from the corn (Zea mays) CDNA libraries SATMON022
BI	CC	and SATMON023. Some of the cdps uniquely identify structural, functional,
BJ	CC	and regulatory genes of corn ear. The polynucleotide sequences are
BK	CC	useful for detecting cdps in a sample, for producing a corn ear-specific
BL	CC	profile of gene transcription, for detecting altered gene expression in
BM	CC	inbred or hybrid plants, and for screening several molecules for specific
BN	CC	binding to the polynucleotide. The cdps are useful to identify, isolate,
BO	CC	or extend identical or related corn-ear nucleic acid sequences from DNA
BP	CC	libraries, and in nucleic acid amplification or hybridization techniques
BQ	CC	to follow the expression of desirable traits through plant breeding
BR	CC	programs. Preferably, the cdps are used to identify, evaluate, alter, or
BS	CC	follow the inheritance of desired characteristics associated with growth
BT	CC	and development, disease resistance, environmental adaptability, quality,
BU	CC	and yield of corn. The cdps are also useful as molecular markers for
BV	CC	studying inheritance and multigene traits in a plant breeding program.
BW	CC	The cdps are useful for producing purified corn ear polypeptides by
BX	CC	recombinant techniques. They are also useful in diagnostic assays to
BY	CC	detect or confirm conditions or diseases associated with abnormal levels
BZ	CC	of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CA	CC	polynucleotides (cdps) of the invention. Note: The sequence data for this
CB	CC	parent did not form part of the printed specification, but was obtained
CC	CC	in electronic format directly from the USPRO web site at
CD	CC	seqdata.uspro.gov/psipsrddentry.html
CE	XX	
CF	SQ	Sequence 270 BP; 83 A; 48 C; 63 G; 76 T; 0 U; 0 Other;
CG	XX	
CH	Query Match	11.3%; Score 225.2; DB 7; Length 270;
CI	Best Local Similarity	96.7%; Pred. No. 5.7e-56;
CJ	Matches 263; Conservative	0; Mismatches 3; Indels 6; Gaps 3
CK		
CL	449 GTTATCCAGCTTTTCGAATGCAGAGAGTGTAATGTGCATAAAAGCCTCATCACAGAG	508
CM	1 GTTATCCAGCTTTTCGAATGCAGAGAGTGTAATGTGCATAAAAGCCTCATCACAGAG	60

QY 509 GTGATGAACCTTGACTCTCCAGANATCTGTAATACAG---CCTTATGAACCTGAGCA 564
 DB 61 GTGATGAACCTTGACTCTCCAG-ATCTGTAATACAGACATGCTAGATGAACCTGAGCA 119
 QY 565 ACTTGAAGGTAATGCTATGATGAAGAGCTGACGCTAGGACCAAGACTACCTGT 624
 DB 120 ACTATGAAGGTAATGCTATGATGAAGAGCTGACGCTAGGACCAAGACTACCTGT 179
 QY 625 TTCAATTGGAAGATGATGTCACACAGATTTTTCCTTATTTGAACAAATTCGTGT 684
 DB 180 TTCAATTGGAAGATGATGTCACACAGATTTTTCCTTATTTGAACAAATTCGTGT 238
 QY 685 CAAGGACGGGAGATGAGGATTTGATTTG 716
 DB 239 CAAGGACGGGAGATGAGGATTTGATTTG 270

RESULT 15

AD08623/c
 ID AD08623 standard; cDNA; 455 BP.

AC AD08623;

DT 04-SEP-2001 (first entry)

DE Oryza sativa aspartate kinase encoding rdrlf.pk005.f20 (EST) cDNA clone.

KM Rice; aspartate kinase; cell transformation; transgenic plant;

XX rdrlf.pk005.f20 clone; ss.

OS Oryza sativa.

XX Key location/Qualifiers
 FT CDS complement (1.342)
 FT /tag= a
 FT /product= "Oryza sativa rdrlf.pk005.f20 (EST) aspartate kinase"

FT W0200146393-A2.

PN 28-JUN-2001.

PD 19-DEC-2000; 2000WO-US034396.

PF 21-DEC-1999; 99US-0172944P.

PR (DUPO) DU PONT DE NEMOURS & CO E I.

PA Falco SC, Famodu OO, Thorpe CJ;

PI WPI; 2001-418057/44.

DR P-PSDB; AAE04355.

XX Novel nucleic acid molecules encoding aspartate kinase useful in cell transformation and transgenic plant production procedures.

XX Claim 8; Page 55; 64pp; English.

XX The present invention relates to nucleic acid molecules encoding aspartate kinase useful in cell transformation and transgenic plant production procedures. Nucleic acid encoding aspartate kinase is useful for producing a transgenic plant, by transforming a plant cell with a chimeric gene comprising nucleic acid encoding aspartate kinase operably linked to a regulatory sequence and regenerating a plant from the transformed plant cell. The present sequence is rdrlf.pk005.f20 (EST)

CC cDNA clone encoding Oryza sativa aspartate kinase

CC Sequence 455 BP; 143 A; 112 C; 74 G; 119 T; 0 U; 7 Other;

XX Query Match 10.3%; Score 204.2; DB 5; Length 455;

XX Best Local Similarity 78.8%; Pred. No. 1.3e-49;

XX Matches 256; Conservative 0; Mismatches 68; Indels 1; Gaps 1;

QY 602 CTTAGACCAAGTACTACTTCTTTCATTTTGAGAAATGATGCTCCACAGATTTTTC- 660
 DB 325 CTTAGACCAAGGATTTACTTGTTCCTTGGTGAATGATGCTTACAGAAATTTGCT 266
 QY 661 TGGTATTGAAACAAATTCGTGCAAGGACGCGAGATGACGATTTGATTTGCTT 720
 DB 265 TGCATTTTGAATTAACCTTGGAAAAAGGCTGGAGATGATGATTTGATTTGGCTT 206
 QY 721 CATTACACTGATGATTTGTAATGCGGATTCCTTAAGAACCTATCTGCTGTTC 780
 DB 205 TATTAAGTCTGATGATTTTACAAATGCGACATTTCTTAAGCGACATATCTGCTGTTC 146
 QY 781 GAAGAGACTTCATGCGGACTGATACAGGATCAGACGATACCTGTTGTTACTGCGTTCCT 840
 DB 145 AAGAGGTTACAGGATGATGATTTGACGACCTGCTATTTCTTAAGTACTGCTTTCCT 86
 QY 841 TGGAAAGGCTGGAATTCGTGCTGTAATCTTTAGGCGGAGTGTGATGACTTGAC 900
 DB 85 TGGAAAAAGATGGAATTCATGCTGTGTCACACCTTTAGGCAAGGTGGACGACTTGAC 26
 QY 901 TGCTACACCATTTGATTAAGCCTTG 925
 DB 25 TGCTACACCATTTGCAAGCGCTTG 1

Search completed: March 23, 2004, 07:25:12
 Job time : 803 secs

GenCore Version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 07:06:09 ; Search time 151 Seconds
(without alignments)
7295.215 Million cell updates/sec

Title: US-09-890-813-5
Perfect score: 1985
Sequence: 1 gcaccagacagacagacagag5.....aataacattcaatctc 1985

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	244.2	12.3	282	4	US-09-313-294A-2808
2	225.2	11.3	270	4	US-09-313-294A-2853
3	129.8	6.5	1664976	4	US-08-915-421B-1
4	86	4.3	640681	4	US-09-790-988-1
5	81.6	4.1	1383	4	US-09-543-681A-3076
6	80	4.0	2589	4	US-09-543-681A-2143
7	77.8	3.9	2915	1	US-07-746-705A-16
8	77.8	3.9	2915	2	US-08-380-182-18
9	69.8	3.5	3412	4	US-09-221-017B-1020
10	68.8	3.5	1437	4	US-09-107-532A-2436
11	66.6	3.4	1830121	4	US-09-557-884-1
12	66.6	3.4	1830121	4	US-09-643-990A-1
13	64.2	3.2	1263	1	US-08-532-828B-11
14	64.2	3.2	1263	1	US-08-532-828B-12
15	64.2	3.2	1643	1	US-08-532-828B-1
16	64.2	3.2	1643	1	US-08-532-828B-2
17	64.2	3.2	1643	1	US-08-532-828B-7
18	64.2	3.2	1643	1	US-08-532-828B-8
19	64.2	3.2	1643	1	US-08-532-828B-9
20	64.2	3.2	1643	1	US-08-532-828B-10
21	64.2	3.2	1643	1	US-08-700-359-7
22	64.2	3.2	1643	1	US-08-700-359-8
23	64.2	3.2	1643	1	US-08-700-359-10
24	64.2	3.2	1643	1	US-08-700-359-11
25	64.2	3.2	1643	1	US-08-700-359-12
26	64.2	3.2	1643	1	US-08-700-359-13
27	64.2	3.2	1643	1	US-08-700-359-14

28	64.2	3.2	1643	2	US-08-967-104-7	Sequence 7, Appli
29	64.2	3.2	1643	3	US-08-985-908-3	Sequence 3, Appli
30	64.2	3.2	1643	3	US-08-985-908-4	Sequence 4, Appli
31	64.2	3.2	1643	3	US-08-985-908-6	Sequence 6, Appli
32	64.2	3.2	1643	3	US-08-852-730-12	Sequence 12, Appli
33	64.2	3.2	1643	3	US-08-852-730-13	Sequence 13, Appli
34	64.2	3.2	1643	3	US-08-852-730-15	Sequence 15, Appli
35	64.2	3.2	1643	3	US-08-985-916-3	Sequence 3, Appli
36	64.2	3.2	1643	3	US-08-985-916-4	Sequence 4, Appli
37	64.2	3.2	1643	3	US-08-985-916-6	Sequence 6, Appli
38	64.2	3.2	1643	3	US-08-985-916-6	Sequence 11, Appli
39	63.6	3.2	19988	4	US-09-526-002-10	Sequence 10, Appli
40	60.4	3.0	771	4	US-08-956-171E-1071	Sequence 1071, Ap
41	59.2	3.0	1317	4	US-09-328-352-2385	Sequence 2385, Ap
42	56.8	2.9	1230025	4	US-09-198-452A-1	Sequence 1, Appli
43	55.8	2.8	1052	4	US-08-961-527-244	Sequence 244, App
44	54.2	2.7	2508	4	US-09-489-039A-1569	Sequence 1569, Ap
45	53.6	2.7	1362	4	US-09-489-039A-1440	Sequence 1440, Ap

ALIGNMENTS

RESULT 1
US-09-313-294A-2808
Sequence 2808, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Lalgudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYTRICLORIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 2808
LENGTH: 282
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700553291H1
US-09-313-294A-2808

Query Match 12.3%; Score 244.2; DB 4; Length 282;
Best Local Similarity 97.1%; Pred. No. 9.4e-63;
Matches 271; Conservative 0; Mismatches 3; Indels 5; Gaps 2;

449 GTTATCATGTTCTGTAATCGAAGAGTGAATATGTCAAAAGCCTACATCAAGACG 508
1 GTTATCATGTTCTGTAATCGAAGAGTGAATATGTCAAAAGCCTACATCAAGACG 60
509 GTTATCATGTTCTGTAATCGAAGAGTGAATATGTCAAAAGCCTACATCAAGACG 564
61 GTTATCATGTTCTGTAATCGAAGAGTGAATATGTCAAAAGCCTACATCAAGACG 119
565 ACATATGAAGTATCGCTATGATGAAGAGTGAATATGTCAAAAGCCTACATCAAGACG 624
120 ACATATGAAGTATCGCTATGATGAAGAGTGAATATGTCAAAAGCCTACATCAAGACG 179
625 TTCTATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 684
180 TTCTATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 239
685 CAAAGCAGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 723
240 CAAAGCAGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 278

RESULT 2
US-09-313-294A-2853

Sequence 2853, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Ialjudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley X.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 2853
LENGTH: 270
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700553382H1
US-09-313-294A-2853

Query Match 11.3%; Score 225.2; DB 4; Length 270;
Best Local Similarity 96.7%; Pred. No. 4.2e-57;
Matches 263; Conservative 0; Mismatches 3; Indels 6; Gaps 3;

QY 449 GTATCCATGTTCTGTAATCGAAGTGGATATGTCAAAAGCCTACATATCAAGACG 508
Db 1 GTATCCATGTTCTGTAATCGAAGTGGATATGTCAAAAGCCTACATATCAAGACG 60
QY 509 GTGATGCACTTGACTTCCAAATCTGTATACAG---CCTTATGAACGTGACGA 564
Db 61 GTGATGCACTTGACTTCCAAATCTGTATACAG---CCTTATGAACGTGACGA 119
QY 565 ACTATTGAAGGTATGCGCTATGATGAAGAGCTGACGCTAGACCAAGTACTACCTGT 624
Db 120 ACTATTGAAGGTATGCGCTATGATGAAGAGCTGACGCTAGACCAAGTACTACCTGT 179
QY 625 TTCAATTGAGAAATGATGTCACCAAGATTTTCTGCTTATTGACAAAATTGCTGT 684
Db 180 TTCAATTGAGAAATGATGTCACCAAGATTTTCTGCTTATTGACAAAATTGCTGT 238
QY 685 CAAAGCAAGGAGATGACGCAATTTGATTG 716
Db 239 CAAAGCAAGGAGATGACGCAATTTGATTG 270

RESULT 3
US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Bull et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6503729
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIORITY FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g

```
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
```

```
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

```
Query Match 6.5% Score 129.8; DB 4; Length 1664976;
Best Local Similarity 49.2%; Pred. No. 2e-26;
Matches 440; Conservative 0; Mismatches 434; Indels 21; Gaps 3;
```

```
389 GCCATGGGAAACCCACCAACACCTTCCTCTCTGCGAGAGAAAGCAGTGGTGA 448
663821 GCATTGGTGGAGATCTCCACCAAGCTTTAGATGATTCGCAAAAGTAGAGAT 963762
449 GTTATCCATGTTTCGAAATCGAAGAGTGAATGCTCAAAAGCCATCATCAAGACG 508
963761 TTTATTAATTTATTAGAGAAACACTCAAACTTATGAAGAGCTATMAAATCGAA 963702
509 GTGATGAACCTTGACCTCCAGNAATCTGTAATCAAGCCTTATGAACTGAGCACTA 568
963701 GAATTAAGAAGAGTAAATAATTAATGACAGCAGATTGAGATTAAGAAGGTT 963642
569 TTGAAGGTATCGCTATGATGAAGAGCTGACCGCTGACCGAGCTACCTGTTTCA 628
963641 TTAAATGGTGTAGCACTTATGGGAGCTTACCAAGTCMAAGCTATATATATCA 963582
629 TTGAGAAATGCATCTCCACAGATTTTTCGCTTATTTGAACAAATTCGTGCAAG 688
963581 TTGAGAGAGGTTCTCTCCACAAATTAAGTAGACTTATGAGATTTAGAGAAAG 963522
689 GCACGCGATATGACGATTTGATATGTTTCACTTCAACTGATGATTTGGTAATGC 748
963521 TCTATTTGCTTGAAGAGGAGAAAGCAGAAATTAACGATTAACACTTTGAAAGTCA 963462
749 GATATCTTAGAAGCAACTATCTGCTGTCGAAGAGCTTCAAGGAGCTGGAATCAG 808
963461 AGAGTTAAAGATTAGAGTTAAAGAGATTTGTACCA-----TATTTAAA 963414
809 GATCCAGGATACCTGTGTTACTTGGGTTCTCTGGAAAGGCTGGAATCTGTGCTGA 868
963413 GAGGCAATTAATCCAGTGTGTTACAGATTTTATAGAA---CCACTGAAGGGTTATATA 963357
869 ACTATTTAGCCGCGAGTGTAGTGTGACTGCTCAACCATTTGTAAGCCTTGGGA 928
963356 ACACCTTAGAAGGCGGAGTGTATTAACAGCCCTTAAATTTGTTAGCTTATGAT 963297
929 CTGAGGAATTCAGATATGAAAGATGTTGATGTTACTTACTTGTATCCAAATATC 988
963296 GCAGATATTAATTAATTTGACAGATGTTTCTGAGTTTATCAACAGACCCAAAGTTA 963237
989 TACCCACATGCAAAACACTGTTCCATCTTAACATTGAAAGGCGACAGAACTTGCTAT 1048
963236 GTTCTACTGCAAGAAGATTTCCAAACTGTTCATATAGAGGTTATGACATAC 963177
1049 TTTGTGTCAGGTTTTCATCAACATCGATGAGACCTGTAGAGAAGGTATATCCA 1108
963176 TTTGAGCTTAAGTTTTCATCAACATTAAGGCTATGAGAAAGGCAATTTCA 963117
1109 GTTAGGTTAAGATTTATCAACCTTAAGCTCCAGGACCCCTTATTAACAGCAAGA 1168
963116 ATATGTTAAGAAATTAATTTGAGCAGAGTGAAGAACTTTAATTAACCAAGATATG 963057
1169 GACATGATAANGCTGCTGTTGTACTAATCAATGATGCTCAAGTCAATGCTCATATG 1228
963056 GAGATGAGTAGATCATGTT-----AAAGTATATCAACAATAAAATATGCTTTA 963003
1229 TTGACATTTGAGACACTGAGATCTTGTGATGATATGTTTTCGGCAAGGCTAT 1283
963002 ATTAACATTTTGGAGCTGGAGTGTGTGAGTTATGATGAGACAGGCGCAAGGATAT 962948
```

```
RESULT 4
US-09-790-988-1/c
```

```

; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGEMOBU, SHUJI
; APPLICANT: MATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHITUKU
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

```

```

Query Match      4.3%; Score 86; DB 4; Length 640681;
Best Local Similarity 48.5%; Pred. No. 1.2e-13;
Matches 236; Conservative 0; Mismatches 251; Indels 0; Gaps 0;

```

```

QY 699 ATGACGCATTGATTTGGTTTCATTACACTGATGAAATTTGGTAATCGGATATCTTAG 758
DB 211266 ATGTACTATTATTATCTCTGTTAAAAATTAGTTGCTATAGTGATATATCTTGATT 211207
QY 759 AAGCAACCTATCTCTGTTGGGAAGAGACTTCATGGGAGCTGATACAGATTCAGCGA 818
DB 211206 CAACGTGATATATCTAATCTAAAAAATATTCAGACATATAATTAATCAAGTA 211147
QY 819 TACCTGTGTTACTGCGGTTCCCTGGGAAGGCGTGGAAATCTGTCGTACTACTTAA 878
DB 211146 ATATTATTATTAATGCTGCTGTTTATTGCGGAAATTAAGTAAAAATTTAGTGTA 211087
QY 879 GCGGAGGTGTAGTACTGACTGCTGATACACCATTTGTAAGCCCTGGGACTGAGAGAA 938
DB 211086 GACGAAATGCTCTGATATTACAGCTGCGGTGAGCGGCTTGTGATGCGAATGTT 211027
QY 939 TTCAGGTATGGAAGAATGTGATGCTGACTTACTTGATGATCCAAATATCAOCCCATG 998
DB 211026 GTGAATCTGACTGATGTCGAGGATTTTACTTCTGATCCAGAAAGTTCCAAATG 210967
QY 999 CAAGACTGTTCCACTACTTAACATTTGAGAGGCGACAACTTGCTTATTTGGTCTC 1058
DB 210966 CTCGTTATTAATAATCAATATCATATCAAGAGCATGSACTGCTTATTTGCTCTA 210907
QY 1059 AGGTTTGCATCCAAATCGATGAGACTGCTAGAGAAAGGATATCCAGTTAGGTTA 1118
DB 210906 AGGTCTACACCTCTGATCTATGAAACCAATGCTCAATTTAAATTCAGTTTAATTA 210847
QY 1119 AGAATTCATACACCTTAAGCTCCAGGCACTTATTCCAGCAAAAGAGACATGATA 1178
DB 210846 AAAATACAAATATATGTTAAATCTATATGAAATTAATTTGGAACAAAGCTTCCAAA 210787
QY 1179 AAGGTCT 1185
DB 210786 AAGATTT 210780

```

```

RESULT 5
US-09-543-681A-3076
; Sequence 3076, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05

```

```

; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3076
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3076

```

```

Query Match      4.1%; Score 81.6; DB 4; Length 1383;
Best Local Similarity 54.3%; Pred. No. 6.3e-14;
Matches 165; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

```

```

QY 869 ACTACTTAAAGCCGAGGTGTAGTGTGACTGCTGATACCAACCTTGTGAAGCTTGAGGA 928
DB 616 ACCACTTAGGGCTGCGGAGATGATTATCTGCGCACTGATTTGTTAAGCATTTAGGG 675
QY 929 CTGAGAGAAATTCAGTATGAGAAAGATGTGATGTTACTTAATTGATCCAAATATC 988
DB 676 ATGTACGTGCTATTTGAGACTGATGCTGCTGATTTATGACACTGATCCAGAAATA 735
QY 989 TACCCATATGAAAGATGTTCCACTTAACTTGAAGAGCCACAGAACTTGCTTAT 1048
DB 736 GTCCCGAAGCAATCGTATGATCATATGCTTTTGAAGAGCGGCTAAATGCAAC 795
QY 1049 TTTGATGCTCAGTTTGCATCCACATCGATGAGACTGCTAGAGAGGTATATTC 1108
DB 796 TTGTTGTATTAATTTCCATCCCGCAACTTATTACTGCACTGCTGATGAGATCCCT 855
QY 1109 GTTAGGTTAAGATTTATACAACTTAAGCTTCCAGGCACTTATTCAGACAAAGA 1168
DB 856 GTGTTTGCCTCAGCAAGCAAGCAAGCAAGGATGTCACATTTGTTGACAAACT 915
QY 1169 GACA 1172
DB 916 GATA 919

```

```

RESULT 6
US-09-543-681A-2143
; Sequence 2143, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2143
; LENGTH: 2589
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2143

```

```

Query Match      4.0%; Score 80; DB 4; Length 2589;
Best Local Similarity 60.6%; Pred. No. 2.8e-13;
Matches 131; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```

```

QY 940 TCAGTATGAAAGATGTGATGTTGTTACTTGTGATCCAAATATCTACCCATGTC 999
DB 738 TGAATTTGACGAGATTTGATGTTTACCTGTGATCCAGCTGCTGCCAGATGC 857
QY 1000 AAAGCTGTTCCATCTTAACATTTGAAGGCGCACAGAACTTGCTTATTTGGTCTA 1059
DB 858 ACACCTATTAAGGATGTCCTATCAAGAGGAGGAGGAACTTCTTACTTTGGTGCCAA 917
QY 1060 GATTTGATCCAAATCGATGAGACTGCTAGAGAGGATATTCAGATTAGGGTTAA 1119
DB 918 AGTCTTCATCAGCAGCATATGCACTATTTGCTCAATTTCAATCCCTTTTATTTAA 977

```

QY 1120 GAATCATACAACTTAAGCTCCAGGACCCCTTAT 1155
DB 978 AATTAGGGGCAATCCAGAGCACCGGTACGTTAT 1013

RESULT 7
US-07-746-705A-16
; Sequence 16, Application US/07746705A
; Patent No. 5451516
; GENERAL INFORMATION:
; APPLICANT: Matthews, Benjamin F.
; APPLICANT: Weisemann, Jane M.
; TITLE OF INVENTION: A Recombinant DNA Molecule Encoding
; TITLE OF INVENTION: A Bifunctional Plant Enzyme: Aspartokinase and Homoserine
; TITLE OF INVENTION: Dehydrogenase
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janelle S. Graeter
; STREET: Bldg. 005, Room 402, BARC-W
; CITY: Beltsville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20705
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07746,705A
; FILING DATE: 19910816
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Graeter, Janelle S.
; REGISTRATION NUMBER: 35,024
; REFERENCE/DOCKET NUMBER: 4000.91
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301)504-5676
; TELEFAX: (301)504-5060
; INFORMATION FOR SEQ. ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2915 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..2593
; US-07-746-705A-16

Query Match 3.9%; Score 77.8; DB 1; Length 2915;
Best Local Similarity 53.9%; Pred. No. 1.3e-12;
Matches 160; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
QY 869 ACTACTTTAGGCCGAGGTGTAGTACTGCTACCAACCATTTGTAAGCCTTGGA 928
DB 695 ACACTTTGAAAAGAGAGCGAAGTACTTTCTGCCCGCTATATGGGTCTTATTAAG 754
QY 929 CTGAGGAATTCAGGTATGAAAGATGTGATGTGTAATCTTACTTGTGATCCAAATATC 988
DB 755 GCTGTCAAGTCAAGATTGGACTGATGTATGTATGTATATAGTCAGATCTCGAAAA 814
QY 989 TACCACATGCAAGAAGCTGTCCATCTTAACATTTGAAGAGCCACAGAACTTGCTTAT 1048
DB 815 GTTAGTAGAGCTGTGTATTAAGACATGTCTTATCAAGAAAGCTGGAGATGTCTAT 874
QY 1049 TTGTGTCTCAGGTTTGTGATCCACATGATGAGAGCTGCTGAGAAAGGTGATATCCA 1108
DB 875 TTGGGGCTTAATGTGTATCATCCCGTACTATCTTCGTGTAGTCGATATGACATTTCA 934

QY 1109 GTTAGGTTAAGATATTCATACACCTTAAGCTCCAGGACCCCTTATTAACAGAA 1165
DB 935 ATTGTAAATAGAAAATATATATCAACTATCTCTCCGGAGCAATGATATGCGAGAA 991

RESULT 8
US-08-380-182-18
; Sequence 18, Application US/08380182
; Patent No. 5858749
; GENERAL INFORMATION:
; APPLICANT: Matthews, Benjamin F.
; APPLICANT: Weisemann, Jane M.
; TITLE OF INVENTION: A Bifunctional Protein From Carrots
; TITLE OF INVENTION: (Daucus carota) with Aspartokinase and Homoserine
; TITLE OF INVENTION: Dehydrogenase Activities
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janelle S. Graeter
; STREET: Room 411, Bldg. 005, BARC-W
; CITY: Beltsville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20705
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/380,182
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Graeter, Janelle S.
; REGISTRATION NUMBER: 35,024
; REFERENCE/DOCKET NUMBER: 0226.94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-504-6629
; TELEFAX: 301-504-5060
; INFORMATION FOR SEQ. ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2915 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Daucus carota
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..2593
; US-08-380-182-18

Query Match 3.9%; Score 77.8; DB 2; Length 2915;
Best Local Similarity 53.9%; Pred. No. 1.3e-12;
Matches 160; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
QY 869 ACTACTTTAGGCCGAGGTGTAGTACTGCTCTCAACCATTTGTAAGCCTTGGA 928
DB 695 ACACTTTGAAAAGAGAGAGTACTTTCTGCCCGCTATATGGGTCTTATTAAG 754
QY 929 CTGAGGAATTCAGGTATGAAAGATGTGATGTGTAATCTTACTTGTGATCCAAATATC 988
DB 755 GCTGTCAAGTCAAGATTGGACTGATGTATGTATGTATATAGTCAGATCTCGAAAA 814
QY 989 TACCACATGCAAGAAGCTGTCCATCTTAACATTTGAAGAGCCACAGAACTTGCTTAT 1048
DB 815 GTTAGTAGAGCTGTGTATTAAGACATGTCTTATCAAGAAAGCTGGAGATGTCTAT 874
QY 1049 TTGTGTCTCAGGTTTGTGATCCACATGATGAGAGCTGCTAGAGAAAGGTGATATCCA 1108

Db 875 TTGGGGGCTAATGTTGTTACATCCCGTACTATCATCTCTGTGATGCGATATGATTCACA 934
Qy 1109 GTTAGGGTTAGATTCATACACCCCTTAAGCTCCAGGACCCCTTATTACAGACAA 1165
Db 935 ATTGATATAGAAATATATTCAACCTATCTGTCGGGAAACATGATATGCGAGAA 991

RESULT 9
US-09-221-017B-1020
Sequence 1020, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & ROEBSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P11546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P12911
FILING DATE: 09-APR-1998
APPLICATION DATA:
PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: McIntoy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ. ID NO: 1020:
SEQUENCE CHARACTERISTICS:
LENGTH: 3412 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc feature
LOCATION: 1...3412
US-09-221-017B-1020

Query Match 3.5%; Score 69.8; DB 4; Length 3412;
Best Local Similarity 52.2%; Pred. No. 3.5e-10;
Matches 155; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

881 CGAGTGTGAGTACTGCTACACCATTTGTAAGCCTTGGAGTGGAGAAATT 940

Db 2797 CGTGGCGGCACGACCTACAGGGCCTCTGTGATGCGAGCAGACCTGATGTGTAATTT 2856
Qy 941 CAGGTATGGAAGATGTTGATGTGTACTTGTGATCTTGAATCAATATCAACCAATGCA 1000
Db 2857 CAGATATGAGGATATGAGCGGATGACACAAATGATCTGTGCTAGACGTTACG 2916
Qy 1001 AAGACTGTTCATCTTAATCAATTTGAAGAGGCACAGAACTTCTATTGTTGCTCAG 1060
Db 2917 GCACCCGTTGGCCGCTTCCATTTTGAAGAAAGCTGCGGAATGGCACACTTCGAGCCAA 2976
Qy 1061 GTTTGATCCACATGATAGACCTGTCTAGAGAGGTATTTCCAGTTAGGTTAAG 1120
Db 2977 ATCCCTCATCCGACCTGTATCAACCGCCGCAAGAGAAATATTCGGTACGTTTACTC 3036
Qy 1121 AATCATACAACTTAAGCTCCAGGACCCCTTATTACAGACAAAGACATGAT 1177
Db 3037 AATACATGATCCCGATGACCCGGGACGCTGATTTCACTGATACCGAAAGAT 3093

RESULT 10
US-09-107-532A-2436
Sequence 2436, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2436:
SEQUENCE CHARACTERISTICS:
LENGTH: 1437 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: ENTEROCOCCUS FAECIUM
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1437
SEQUENCE DESCRIPTION: SEQ ID NO: 2436:
US-09-107-532A-2436

Query Match 3.5%; Score 68.8; DB 4; Length 1437;
 Best Local Similarity 52.6%; Pred. No. 4.2e-10;
 Matches 113; Conservative 0; Mismatches 153; Indels 3; Gaps 1;

QY 875 TTAGGCGAGGTGGTGTGATGACTGTACACCATGTGTAAGCCTGGAGCTGAGA 934
 DB 628 TTGGGCGCGTGGGGAAGCAGCATATCCGCGGATCTGCGCAAAATGTTCCAGCGGA 687
 QY 935 GAAATTCAGTATGAAAGATGTTGATGTTACTTACTTGTATCCAAATCTTACCCA 994
 DB 688 AGTTATGAGATTTGACGATGTTTTCAGAAATCATGATGCTGATCCCGAATCATCAT 747
 QY 995 CATGCAAGCTGTTCATACATTAACATTTGAAGGCGACAGACTGCTTATTTGGT 1054
 DB 748 CATCTTAAATATCATGAGGAATCTTTTAAAGAACTGAGAGATGCTTATATGAGG 807
 QY 1055 GCTCAGGTTTTCATCCACATGATGAGACCTGCTAGAGAGTGAATATTCAGTTAGG 1114
 DB 808 GCTACGTTCTTCATGAGCAGCTGTTTCCAGTTCAAGAGGCTGATATCCCATCCAA 867
 QY 1115 GTTAAAGATTATACAAACCTTAAAGCTCCAGCACCTTATACAGAC---AAAGAGAC 1171
 DB 868 ATAAAAATATCCATATATCTTTCCTCAGAGGACAAAAATATCGAATCAGAAATAGAA 927
 QY 1172 ATGATATANGTCTGCTGTACTACTAG 1200
 DB 928 AAGGAAATGAGACTGACAGAAATAGCTG 956

RESULT 11
 US-09-557-884-1/c
 ; Sequence 1, Application US/09557884
 ; Patent No. 6506581

GENERAL INFORMATION:

APPLICANT: Fleischmann et al.

TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/557,884

FILING DATE: 25-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,102

FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB16PF3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Query Match 3.4%; Score 66.6; DB 4; Length 1830121;
 Best Local Similarity 52.3%; Pred. No. 1.4e-07;
 Matches 147; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 876 TAGGCGAGGTGGTGTGATGACTGTACACCATGTGTAAGCCTGGAGCTGAGAG 935
 DB 95676 TCGGTGATATGCTCAGCATATATCTGCGGCTGTTAGCCGCTGTTTAGCCCAAGTG 95617
 QY 936 AATTCAGTATGAAAGATGTTGATGTTACTTACTTGTATCCAAATCTTACCCAC 995
 DB 95616 TATGCGAAATTTGACCCATGTCAGCGCGTTTACTTGGATCCCGGTTAGTGCCTG 95557
 QY 996 ATGCAAGCTGTTCATACATTAACATTTGAAGGCGACAGACTGCTTATTTGGTG 1055
 DB 95556 ATGCAAGCTTATTAACCAACGCTTTTATCGTGAAGCATGAACTCTTATTTCCGTTG 95497
 QY 1056 CTCAGGTTTTCATCCACATGATGAGACCTGCTAGAGAGTGAATATTCAGTTAGG 1115
 DB 95496 CGAAAGTATTCATCTCTGATGATGAGGCGATTTGCCCAAAATATCTTGTGTCA 95437
 QY 1116 TTAAGAAATCATACAAACCTTAAAGCTCCAGGACCCCTATT 1156
 DB 95436 TTAATAATACAGAAATCTTCTGACCCAGGTTGATTAT 95396

RESULT 12
 US-09-643-990A-1/c
 ; Sequence 1, Application US/09643990A
 ; Patent No. 6528289

GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann

Mark D. Adams

Owen White

Hamilton O. Smith

J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/643,990A

FILING DATE: 23-Aug-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,429

FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787

FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB186P1C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-610-5790

TELEFAX: 301-610-5790

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match 3.4%; Score 66.6; DB 4; Length 1830121;
Best Local Similarity 52.3%; Pred. No. 1.4e-07;
Matches 147; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 876 TAGGCGAGGTGTAGTACTTGAAGTCAACCATTTGTAAGCCTGGGACTGAG 935
DB TCGGCTGTAATGGCTCAACACTATTCTGCGCTCTTTAGCCGCTGTTTGGCGAAGTG 95617
QY 936 AATTCAGGTATGGAAGATGTGATGTGTACTTGTGATCCAAATATCTAACCA 995
DB 95616 TATGCGAATTTGGACGATGGACGGCGTTTATCTTGGAATCCGCGTTTGTGCTG 95557
QY 996 ATGCAAGACTGTCTCACTTAACTTGAAGGCGCAGAACTGCTTATTTGGTG 1055
DB 95556 ATGCAAGCTTTATTAACCAAGCTTTCTTATCTGTGAAGCGATGGAACCTCTTATTTGGTG 95497
QY 1056 CTCAGGTTTGCATCCACATCGATGAGACCTGCTAGAGAAGGTGATATCCAGTTAGGG 1115
DB 95496 CGAAGTGTATCATCTCTGTGAGATGGGCCATTGTTGCCAATAATTTCTTGTGTCA 95437
QY 1116 TTAGAATTCATACACCTTAAGCTCCAGGACCTTATT 1156
DB 95436 TTTAAATATACAGAAATCCTTCTGACACGATTTGATT 95396

RESULT 13
US-08-532-828B-11
Sequence 11, Application US/08532828B

PATENT INFORMATION:
PATENT NO. 5688671
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: OGAWA, Yuki
APPLICANT: SUZUKI, Tomoko
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
TITLE OF INVENTION: MUTANT ASPARTOKINASE GENE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS TEXT EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532, 828B
FILING DATE: 27-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: JP 5-101450
FILING DATE: 27-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-764-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1263 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC13869
US-08-532-828B-11

Query Match 3.2%; Score 64.2; DB 1; Length 1263;
Best Local Similarity 49.1%; Pred. No. 9e-09;
Matches 199; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 757 AGAAGCACTATCTCTGTGTGCGAAGACTTCATGCGGACTGATACAGATCCAGC 816
DB 327 AAAGCAGCATTTGTAAGCTCACACCGGCTGTGTGCGTAAGCACTGATGAGGCAA 386
QY 817 GATACCTGTGTACTAGGCGTCTCTTGGGAAGGCTGGAATGTGCTGTACTATT 876
DB 387 GATCTGATTTGTCTGTGTTTCAAGGTGTTAATAAGAAACCGCATGTACACGTT 446
QY 877 AGCGGAGGTGTAGTACTGACTGCTCAACCACTTGTAAGCCTTGGACTGAGA 936
DB 447 GGTGTGTGTGTGTGTGACACCACTGACGTGGGTGTGCAAGCTGTTGAAGCTGATGT 506
QY 937 AATTCAGTATGGAAGATGTGATGTGTACTTGTGTATCCAAATATCTACCCACA 996
DB 507 GTGTGATTTACTCTGACGCTTACGCTGTGTATACCGCTGACCGCGCATGTTCTTA 566
QY 997 TGCAAGACTGTCTCACTTAACTTGAAGGCGCAGAACTGCTTATTTGGTG 1056
DB 567 TGCAAGACTGTGAAAGCTCAAGCTTGAAGAAATGCTGAACTGTGCTGTTGGCTC 626
QY 1057 TCAGGTTTGCATCCACATCGATGAGACCTGCTAGAGAAGGTATATCCAGTTAGGT 1116
DB 627 CAAGATTTGTGTGTGCGAGTGTGAATAGCTCTGTGATTCATATGTGCCACTTGGCGT 686
QY 1117 TAAGATTCATACACCTTAAGCTCCAGGACCTTATTACAG 1161
DB 687 ACGCTGCTTATTA---GTAAATATCCGCGACTTGTGATTTGGCGG 728

RESULT 14
US-08-532-828B-12
Sequence 12, Application US/08532828B

PATENT INFORMATION:
PATENT NO. 5688671
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: OGAWA, Yuki
APPLICANT: SUZUKI, Tomoko
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
TITLE OF INVENTION: MUTANT ASPARTOKINASE GENE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS TEXT EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532, 828B
FILING DATE: 27-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: JP 5-101450
FILING DATE: 27-APR-1993
ATTORNEY/AGENT INFORMATION:

NAME: NORMAN F. OBLON
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 10-764-0 PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1263 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: genomic DNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Corynebacterium glutamicum
 STRAIN: AJ3463
 US-08-532-828B-12

Query Match 3.2% Score 64.2; DB 1; Length 1263;
 Best Local Similarity 49.1%; Pred. No. 9e-09;
 Matches 199; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 757 AGAAGCAACCTATCTCTGCTGTTGCGAAGAGACTTCATGGGACTGATACAGATCCAGC 816
 DB 327 AAACGACACGATGTTGACGTACACCGGGTGTGTGCGTAGACACTCGATGAGGGCAA 386
 QY 817 GATACCTGTGTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876
 DB 387 GATCTGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 446
 QY 877 AGCGCGAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 936
 DB 447 GGGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 506
 QY 937 AATTCAGTATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 996
 DB 507 GTGTGATTTACTGCGAGCTGACGCTGTGTATACCCCTGACCCGCGCATGCTTCTAA 566
 QY 997 TCGAAGACTGTTCATCTTACTTAACTTGAAGAGCCACAGAACTTGTCTTATTTTGTGTC 1056
 DB 567 TCGACAGAGAGCTGGAAGAGCTCAGCTTGAAGAAATGCTGGAATGCTGTGCTGCTGCTC 626
 QY 1057 TCAGGTTTGCATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1116
 DB 627 CAAGATTTGTGCTGCGCAGGTGTGAATACGCTGTCATTCATGTGCGACTTGCCTG 686
 QY 1117 TAAGATTCATCAACCTTAAGCTCCAGGCAACCTTATTACAG 1161
 DB 687 AGCGTGTCTTATA---GTAAATGATCCCGGCACTTTGATTGCCG 728

RESULT 15
 US-08-532-828B-1
 ; Sequence 1, Application US/08532828B
 ; Patent No. 5688671
 ; GENERAL INFORMATION:
 ; APPLICANT: SUGIMOTO, Masakazu
 ; APPLICANT: OGAWA, Yuki
 ; APPLICANT: SUZUKI, Tomoko
 ; APPLICANT: TANAKA, Akiko
 ; APPLICANT: MATSUI, Hiroshi
 ; TITLE OF INVENTION: MUTANT ASPARTOKINASE GENE
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT, P.C.
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 ; CITY: ARLINGTON
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: MS-DOS TEXT EDITOR
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/532,828B
 FILING DATE: 27-OCT-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 5-101450
 FILING DATE: 27-APR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: NORMAN F. OBLON
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 10-764-0 PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1643 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: genomic DNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Corynebacterium glutamicum
 STRAIN: ATCC 13869
 US-08-532-828B-1

Query Match 3.2% Score 64.2; DB 1; Length 1643;
 Best Local Similarity 49.1%; Pred. No. 1.1e-08;
 Matches 199; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 757 AGAAGCAACCTATCTCTGCTGTTGCGAAGAGACTTCATGGGACTGATACAGATCCAGC 816
 DB 543 AAACGACACGATGTTGACGTACACCGGGTGTGTGCGTAGACACTCGATGAGGGCAA 602
 QY 817 GATACCTGTGTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876
 DB 603 GATCTGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 662
 QY 877 AGCGCGAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 936
 DB 663 GGGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722
 QY 937 AATTCAGTATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 996
 DB 723 GTGTGATTTACTGCGAGCTGTGTATACCCCTGACCCGCGCATGCTTCTAA 782
 QY 997 TCGAAGACTGTTCATCTTACTTAACTTGAAGAGCCACAGAACTTGTCTTATTTTGTGTC 1056
 DB 783 TCGACAGAGAGCTGGAAGAGCTCAGCTTGAAGAAATGCTGGAATGCTGTGCTGCTGCTC 842
 QY 1057 TCAGGTTTGCATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1116
 DB 843 CAAGATTTGTGCTGCGCAGGTGTGAATACGCTGTCATTCATGTGCGACTTGCCTG 902
 QY 1117 TAAGATTCATCAACCTTAAGCTCCAGGCAACCTTATTACAG 1161
 DB 903 AGCGTGTCTTATA---GTAAATGATCCCGGCACTTTGATTGCCG 944

Search completed: March 23, 2004, 11:07:11
 Job time : 161 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 09:36:45 ; Search time 699 Seconds
(without alignments)
10504.749 Million cell updates/sec

Title: US-09-890-813-5
Perfect score: 1985
Sequence: 1 gcacccagacagcagcagagg.....atatacatcatcatc 1985

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1845576744 residues
Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgnt2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgnt2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgnt2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgnt2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgnt2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgnt2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgnt2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgnt2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgnt2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
- 10: /cgnt2_6/prodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgnt2_6/prodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgnt2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgnt2_6/prodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgnt2_6/prodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgnt2_6/prodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgnt2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
- 17: /cgnt2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
- 18: /cgnt2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1983	99.9	1985	9 US-09-890-813-5	Sequence 5, Appl1
2	1863.2	93.9	1953	9 US-09-890-813-7	Sequence 7, Appl1
3	1416.6	71.4	1533	12 US-10-425-114-5334	Sequence 5934, Ap
4	831.4	41.9	1115	12 US-10-425-114-33565	Sequence 33565, A
5	745.4	37.6	1180	12 US-10-425-114-6176	Sequence 6176, Ap
6	740.6	37.3	2882	12 US-10-424-599-28065	Sequence 28065, A
7	707.6	35.6	1710	9 US-09-938-842A-1744	Sequence 1744, Ap
8	707.6	35.6	1710	11 US-09-938-842A-1744	Sequence 1744, Ap
9	674.6	34.0	1568	9 US-09-890-813-15	Sequence 15, Appl
10	655.4	33.0	1582	12 US-10-425-114-3869	Sequence 3869, Ap
11	608.2	30.6	1361	9 US-10-425-114-8935	Sequence 8935, Ap
12	555	28.0	365	9 US-09-890-813-1	Sequence 1, Appl1
13	533.2	26.9	2490	12 US-10-424-599-28060	Sequence 28060, A
14	515	25.9	1671	12 US-10-425-114-7821	Sequence 7821, Ap
15	470.4	23.7	1260	12 US-10-425-114-7451	Sequence 7451, Ap

16	448.6	22.6	513	9 US-09-890-813-3	Sequence 3, Appl1
17	371.2	18.7	847	9 US-09-890-813-11	Sequence 11, Appl
18	302.2	15.2	944	12 US-10-425-114-9815	Sequence 9815, Ap
19	236.2	11.9	825	12 US-10-425-114-8840	Sequence 8840, Ap
20	204.2	10.3	455	9 US-09-890-813-9	Sequence 9, Appl1
21	164.2	8.3	1060	12 US-10-424-599-28063	Sequence 28063, A
22	158.2	8.0	646	9 US-09-890-813-13	Sequence 13, Appl
23	149.2	7.5	276	9 US-09-294-093B-2684	Sequence 2684, Ap
24	140.8	7.1	282	9 US-09-923-876-5481	Sequence 5481, Ap
25	140.8	7.1	282	11 US-10-369-493-54150	Sequence 54150, Ap
26	129.8	6.5	1422	15 US-10-369-493-34902	Sequence 34902, A
27	129	6.5	1392	15 US-10-369-493-34902	Sequence 34902, A
28	112.2	5.7	286	9 US-09-294-093B-770	Sequence 770, App
29	101.8	5.1	1289	15 US-10-369-493-34655	Sequence 24655, A
30	95.6	4.8	285	9 US-09-294-093B-383	Sequence 383, App
31	86	4.3	640681	9 US-09-790-989-1	Sequence 1, Appl1
32	84	4.2	2751	9 US-09-938-842A-1105	Sequence 1105, Ap
33	84	4.2	2751	11 US-09-938-842A-1105	Sequence 1105, Ap
34	79.6	4.0	2448	12 US-10-282-122A-30697	Sequence 30697, A
35	79.4	4.0	2457	12 US-10-282-122A-32787	Sequence 32787, A
36	79	4.0	1584	15 US-10-369-493-25428	Sequence 25428, A
37	79	4.0	2433	12 US-10-282-122A-12611	Sequence 12611, A
38	77.4	3.9	1444	15 US-10-369-493-24150	Sequence 24150, A
39	77	3.9	1443	15 US-10-369-493-24150	Sequence 24150, A
40	76.8	3.9	1203	12 US-10-282-122A-18184	Sequence 18184, A
41	76.2	3.8	2460	12 US-10-282-122A-41673	Sequence 41673, A
42	76.2	3.8	3141	12 US-10-424-599-92508	Sequence 92508, A
43	76.2	3.8	3263	12 US-10-424-599-92505	Sequence 92505, A
44	76.2	3.8	3561	12 US-10-424-599-92505	Sequence 92505, A
45	75.6	3.8	1467	15 US-10-369-493-46250	Sequence 46250, A

ALIGNMENTS

RESULT 1
US-09-890-813-5
Sequence 5, Appl1
Publication No. US09890813
GENERAL INFORMATION:
APPLICANT: E. I. du Pont de Nemours and Company
TITLE OF INVENTION: Aspartate Kinase
FILE REFERENCE: BS1430 PCT
CURRENT APPLICATION NUMBER: US/09/890, 813
PRIORITY FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 60/172944
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Microsoft Office 97
SEQ ID NO 5
LENGTH: 1985
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (532)
NAME/KEY: unsure
LOCATION: (1180)
US-09-890-813-5

Query Match 99.9%; Score 1983; DB 9; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCACCAAGACAGACAGAGAGACTCAGAAATGCCAATCCAGTGCATGCGCTCCGCCGCC 60
DB 1 GCACCAAGACAGACAGAGAGACTCAGAAATGCCAATCCAGTGCATGCGCTCCGCCGCC 60
QY 61 CCGCCGCGTTCGTCGTCATCCTCCGCGAGCTCTGGACATGTTTCAGAGACATGCGCTG 120
DB 61 CCGCCGCGTTCGTCGTCATCCTCCGCGAGCTCTGGACATGTTTCAGAGACATGCGCTG 120
QY 121 CTTGGATCCGACCGGAGCTCTCGGATGCAAGAGGTTTCATGATGTTGCCGCACTC 180

Db 121 CTTCCGTTACCCGAGCCGAGCTCGCGGTGCAAGAGGTTGTCAAATGGTGGTCCGAGCTC 180
 Qy 181 CACACACCGTCGAGGCAAGCAAGCGGACGCGGAGCGGCTTGGAGGCGCTGTTCT 240
 Db 181 CACACACCGTCGAGGCAAGCAAGCGGACGCGGAGCGGCTTGGAGGCGCTGTTCT 240
 Qy 241 CGAGAGGCTCGAGGATGAGAGGATTTGGGGATCACTCAAGCTGTGATGAAGTTGGGG 300
 Db 241 CGAGAGGCTCGAGGATGAGAGGATTTGGGGATCACTCAAGCTGTGATGAAGTTGGGG 300
 Qy 301 GTCTCGGTGTGTGTGGCGCGGAGATGGCTGAGGTCGCGGCTCATCTGAGCTTCC 360
 Db 301 GTCTCGGTGTGTGTGGCGCGGAGATGGCTGAGGTCGCGGCTCATCTGAGCTTCC 360
 Qy 361 CGAGAGCGCCCGTGTGTCTCTTCCATGGGGAACCAACCAACCTTCTCT 420
 Db 361 CGAGAGCGCCCGTGTGTCTCTTCCATGGGGAACCAACCAACCTTCTCT 420
 Qy 421 TCGTGAAGAGAGGCGAGTGGGTGAGTATCCATGTTCTGAAATCGAAGGTGGA 480
 Db 421 TCGTGAAGAGAGGCGAGTGGGTGAGTATCCATGTTCTGAAATCGAAGGTGGA 480
 Qy 481 TATGTCAAAGCTCATATCAAGACGCTGATGAACCTTGACCTTCCAGNATCTGAA 540
 Db 481 TATGTCAAAGCTCATATCAAGACGCTGATGAACCTTGACCTTCCAGNATCTGAA 540
 Qy 541 TACAGCCTTTATGAATCGAGCACTATTGAAGGTATCGTATGATGAAGAGCTGAC 600
 Db 541 TACAGCCTTTATGAATCGAGCACTATTGAAGGTATCGTATGATGAAGAGCTGAC 600
 Qy 601 GCCTAGAGCAGTGAATCTGTTTCATTTGAGAAATGCAATGTCACAGATTTTTC 660
 Db 601 GCCTAGAGCAGTGAATCTGTTTCATTTGAGAAATGCAATGTCACAGATTTTTC 660
 Qy 661 TCGTTATTTGAACAAATTCGTGTCAAGGACGAGCAAGTATGACGATTTGATTTG 720
 Db 661 TCGTTATTTGAACAAATTCGTGTCAAGGACGAGCAAGTATGACGATTTGATTTG 720
 Qy 721 CATTAACACTGATGATTTGATGATGCGATATCTTAGAGCAACCTATCCGTGTTGC 780
 Db 721 CATTAACACTGATGATTTGATGATGCGATATCTTAGAGCAACCTATCCGTGTTGC 780
 Qy 781 GAAGAGCTTCATGGGAGCTGATACAGATCCAGCATCTGTTGTTACTGGGTCT 840
 Db 781 GAAGAGCTTCATGGGAGCTGATACAGATCCAGCATCTGTTGTTACTGGGTCT 840
 Qy 841 TGGGAGGCTGGAATCTGTGTCTGATCTACTTAGGCGAGGTGATGACTTGC 900
 Db 841 TGGGAGGCTGGAATCTGTGTCTGATCTACTTAGGCGAGGTGATGACTTGC 900
 Qy 901 TCGTCAACACCATTTGTAAGCTTGGGACTGAGAGAAATTCAGGTATGAAAGATTTGA 960
 Db 901 TCGTCAACACCATTTGTAAGCTTGGGACTGAGAGAAATTCAGGTATGAAAGATTTGA 960
 Qy 961 TGGTGTACTTACTTGTATCCAAATATCTACCAACATGCAAGATCTTCACTTAC 1020
 Db 961 TGGTGTACTTACTTGTATCCAAATATCTACCAACATGCAAGATCTTCACTTAC 1020
 Qy 1021 ATTGAAGAGGCAAGAACTTGTATTTGGTGTCAAGTTTGGATTCACATTCAT 1080
 Db 1021 ATTGAAGAGGCAAGAACTTGTATTTGGTGTCAAGTTTGGATTCACATTCAT 1080
 Qy 1081 GAGACCTGTAGAGAGTGTATTCAGTGTAGGTTAAGATTCATCAACCTTAAAG 1140
 Db 1081 GAGACCTGTAGAGAGTGTATTCAGTGTAGGTTAAGATTCATCAACCTTAAAG 1140
 Qy 1141 TCCAGGACCTTATTTACAGACAAAGAGACATGTATANGTCTGGTTGATCAATAG 1200
 Db 1141 TCCAGGACCTTATTTACAGACAAAGAGACATGTATANGTCTGGTTGATCAATAG 1200
 Qy 1201 CATAGTGTCAATGTAATGTCTATGTTGATGATTTGGAGCACTGGAGTGTGTTCA 1260

Db 1201 CATAGTGTCAATGTAATGTCTATGTTGACATTTGTAGACATTTGTAGCACTGGATGCTCA 1260
 Qy 1261 GTATGTTTTCGCAAGAGGATATAGGATTTGTATATTAATGAAGTATGATATCTGT 1320
 Db 1261 GTATGTTTTCGCAAGAGGATATAGGATTTGTATATTAATGAAGTATGATATCTGT 1320
 Qy 1321 GATTTGTGTGTACCAAGTGAAGTATGTTTCTGTGTCTCACTTATCCATCAAGATCTG 1380
 Db 1321 GATTTGTGTGTACCAAGTGAAGTATGTTTCTGTGTCTCACTTATCCATCAAGATCTG 1380
 Qy 1381 GAGTAGGAGACTGATACAGCAGGAGTGAACCTTGACATGTATGTTGAAGAGCTTGA 1440
 Db 1381 GAGTAGGAGACTGATACAGCAGGAGTGAACCTTGACATGTATGTTGAAGAGCTTGA 1440
 Qy 1441 AATAGCATTTGTGTCTTACTTACAGCAGAGGCGATTAATTCATTAATCGGAATGTGA 1500
 Db 1441 AATAGCATTTGTGTCTTACTTACAGCAGAGGCGATTAATTCATTAATCGGAATGTGA 1500
 Qy 1501 GCAATCGTCTGTATACAGAAAGACGAGAGCTGTGCTGAGGAAAGTGGGGTTAATGT 1560
 Db 1501 GCAATCGTCTGTATACAGAAAGACGAGAGCTGTGCTGAGGAAAGTGGGGTTAATGT 1560
 Qy 1561 TCAGATGATCTCGCAAGAGCGTCAAGGTTTACATGTCTGATATGATGATGCGA 1620
 Db 1561 TCAGATGATCTCGCAAGAGCGTCAAGGTTTACATGTCTGATATGATGATGCGA 1620
 Qy 1621 TGCAGAGGACCTGAGAGGCGCTTCAACAGGCGTTCTTGAAGACGATGCTATCACA 1680
 Db 1621 TGCAGAGGACCTGAGAGGCGCTTCAACAGGCGTTCTTGAAGACGATGCTATCACA 1680
 Qy 1681 AGTCAAGCGAGAACTTACTGTGGGCTGATCAAGTACGCTTGTGGTCCAGGCGT 1740
 Db 1681 AGTCAAGCGAGAACTTACTGTGGGCTGATCAAGTACGCTTGTGGTCCAGGCGT 1740
 Qy 1741 GTTATCTGTATGATTTCCCACTGCGCTCCATGAGAGGCAATGGGCAATGGATCATGATC 1800
 Db 1741 GTTATCTGTATGATTTCCCACTGCGCTCCATGAGAGGCAATGGGCAATGGATCATGATC 1800
 Qy 1801 ATGTTTGTCTGAAACAAAGTATGTTTCCAGGTTCTCAAGCAATGATCTGCAAACTGTGT 1860
 Db 1801 ATGTTTGTCTGAAACAAAGTATGTTTCCAGGTTCTCAAGCAATGATCTGCAAACTGTGT 1860
 Qy 1861 TTTCTGTTTGAAGCTGTTGAGACACAGTGAAGCTGAGACCGATTTGTCAACAGAT 1920
 Db 1861 TTTCTGTTTGAAGCTGTTGAGACACAGTGAAGCTGAGACCGATTTGTCAACAGAT 1920
 Qy 1921 GCGAAGCCTGTGATTAATTCACACTGTCTTAATCAATATATTAATAAATTATCA 1980
 Db 1921 GCGAAGCCTGTGATTAATTCACACTGTCTTAATCAATATATTAATAAATTATCA 1980
 Qy 1981 TATCT 1985
 Db 1981 TATCT 1985

RESULT 2
 US-09-890-813-7
 ; Sequence 7, Application US/09890813
 ; Publication No. US20020183486A1
 ; GENERAL INFORMATION:
 ; APPLICANT: E.I. du Pont de Nemours and Company
 ; TITLE OF INVENTION: Aspartate Kinase
 ; FILE REFERENCE: B01430 PCT
 ; CURRENT APPLICATION NUMBER: US/09/890, 813
 ; PRIORITY FILING DATE: 2001-08-02
 ; PRIOR APPLICATION NUMBER: 60/172944
 ; PRIORITY FILING DATE: 1999-12-21
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 7
 ; LENGTH: 1953
 ; TYPE: DNA
 ; ORGANISM: Zea mays

US-09-890-813-7

Query Match 93.9%; Score 1863.2; DB 9; Length 1953;
 Best Local Similarity 98.4%; Pred. No. 0;
 Matches 1941; Conservative 0; Mismatches 8; Indels 23; Gaps 5;

QY 11 GAGCAGAGAGACTCAGAAATGGCAATCCAGTGCATCGGCTGCCGCCGCCGCCGCTC 70
 Db 1 GAGCAGAGAGACTCAGAAATGGCAATCCAGTGCATCGGCTGCCGCCGCCGCCGCTC 60

QY 71 GTTCCGTGCATCTCCGCGAGCTCTGACATGTTTCAAGACTGCGCTGCTTGGTACC 130
 Db 61 GTTCCGTGCATCTCCGCGAGCTCTGACATGTTTCAAGACTGCGCTGCTTGGTACC 120

QY 131 CGAACCGGGGCTCGCGGTGCAAGAGGTTGTCAATGCTGTGCGCGACTCCACAGCCGT 190
 Db 121 CGAACCGGGGCTCGCGGTGCAAGAGGTTGTCAATGCTGTGCGCGACTCCACAGCCGT 180

QY 191 CGGGCCAGCAGACCGAGCGCGGAGCGCGCTCTTGGGGCGCTGTTCTCGAAGGCTC 250
 Db 181 CGGGCCAGCAGACCGAGCGCGGAGCGCGCTCTTGGGGCGCTGTTCTCGAAGGCTC 240

QY 251 GGGATGAGAGGATTTGGGGGATCAGCTCAGCGTGTGATGAATTTGGGGGGTCTCGATG 310
 Db 241 GGGATGAGAGGATTTGGGGGATCAGCTCAGCGTGTGATGAATTTGGGGGGTCTCGATG 300

QY 311 TCGTCGGCGCGAGATGCTGAGTGGCGCGCTCATCTGACGTTCCCGAGAGCGC 370
 Db 301 TCGTCGGCGCGAGATGCTGAGTGGCGCGCTCATCTGACGTTCCCGAGAGCGC 360

QY 371 CCGGTGCTGTTCTCTGCGCATGGGGAACACCAACACCTTCTGCTGGAGAG 430
 Db 361 CCGGTGCTGTTCTCTGCGCATGGGGAACACCAACACCTTCTGCTGGAGAG 420

QY 431 AAGCAGTAGGAGTGTGAGTATTCATGTTCTGAAATGGAAGATGGAATATGCTCAA 490
 Db 421 AAGCAGTAGGAGTGTGAGTATTCATGTTCTGAAATGGAAGATGGAATATGCTCAA 480

QY 491 AGCTTACATATCAAGACGCTGATGAATCTTGAATCTTCAAGATCTTGAATCAAGCTT 550
 Db 481 AGCTTACATATCAAGACGCTGATGAATCTTGAATCTTCAAGATCTTGAATCAAGCTT 539

QY 551 T----ATGAAGTGAAGAACTATTGAAGTATCGTATGATGAAGAGCTGAGCCCTAG 606
 Db 540 GCTAGATGAATGAGCAACTATTGAAGTATCGTATGATGAAGAGCTGAGCCCTAG 599

QY 607 GACCACTGACTACCTTGTTCATTTGAGAAATGATGTCACACAGATTTTTTCTGCTTA 666
 Db 600 GACCACTGACTACCTTGTTCATTTGAGAAATGATGTCACACAGATTTTTTCTGCTTA 659

QY 667 TTGGAACAAATTCGCTGCAAGGACGAGAGTATGACGATTTGATTTGTTGCTATAC 726
 Db 660 TTGGAACAAATTCGCTGCAAGGACGAGAGTATGACGATTTGATTTGTTGCTATAC 719

QY 727 AACTGATGATTTGATATGCGGATATCTTGAAGCAACTATCTGCTGTTGCGAAG 786
 Db 720 AACTGATGATTTGATATGCGGATATCTTGAAGCAACTATCTGCTGTTGCGAAG 779

QY 787 ACTTCATGAGGATCTGATACAGATTCAGCGATCTGTTGTTACTGGGTTCTTGGGAA 846
 Db 780 ACTTCATGAGGATCTGATACAGATTCAGCGATCTGTTGTTACTGGGTTCTTGGGAA 839

QY 847 GGGCTGAAATCTGCTGCTGTAATCTATTAGCGAGGTGATGACTTGAAGCTGCTAC 906
 Db 840 GGGCTGAAATCTGCTGCTGTAATCTATTAGCGAGGTGATGACTTGAAGCTGCTAC 899

QY 907 AACCATTTGTAAGCCTTGGGACTGAGAAATTCAGGTATGGAAGATGTTAGTGT 966
 Db 900 AACCATTTGTAAGCCTTGGGACTGAGAAATTCAGGTATGGAAGATGTTAGTGT 959

QY 967 ACTTACTTGTGATCAAAATATCTACCAACATGCAAGATGTTTCAATCTTAACTTTGA 1026
 Db 960 ACTTACTTGTGATCAAAATATCTACCAACATGCAAGATGTTTCAATCTTAACTTTGA 1019

QY 1027 AAGGCGACAGAACTGCTATTTTGGTGTCAAGTTTGGATCCCAATCGATGAGACC 1086
 Db 1020 AAGGCGACAGAACTGCTATTTTGGTGTCAAGTTTGGATCCCAATCGATGAGACC 1079

QY 1087 TGTAGAGAGGTGATATTCAGTTAGGTTAAGATTCATACACCTTAAAGCTCCAGG 1146
 Db 1080 TGTAGAGAGGTGATATTCAGTTAGGTTAAGATTCATACACCTTAAAGCTCCAGG 1139

QY 1147 CACCTTATTCACAGACAAAGACATGATTAAGGTCTGGTGTACTACTAGATAGT 1206
 Db 1140 CACCTTATTCACAGACAAAGACATGATTAAGGTCTGGTGTACTACTAGATAGT 1193

QY 1207 GCTCAAGTCMAATGTCATATGTTGACATTTGAGACATGAGACTGGATGCTGATGAG 1266
 Db 1194 GCTCAAGTCMAATGTCATATGTTGAGACATTTGAGACATGAGACTGGATGCTGATGAG 1253

QY 1267 TTTTCTGGCAAGGATACAGATTTTGTATATGAAATCTATATCTGTGATG 1326
 Db 1254 TTTTCTGGCAAGGATATTTGTATAT-----TTGAAGATCTATATATCTGTGATG 1307

QY 1327 TGTGCTACAGTGAAGTTAGTGTCTGTCTCATCTTATCATCAAGATCTGAGTAG 1386
 Db 1308 TGTGCTACAGTGAAGTTAGTGTCTGTCTCATCTTATCATCAAGATCTGAGTAG 1367

QY 1387 GGAATGATACAGCAGGCAAGTGAATTGACCAATGTAGTGAAGACTTGAGAAATAGC 1446
 Db 1368 GGAATGATACAGCAGG-----GAATGACCAATGTAGTGAAGACTTGAGAAATAGC 1421

QY 1447 AATGTGTGCTACTTACAGACAGAGGCGCATATTTCACTTATCGAAATGTGAGCAATC 1506
 Db 1422 AATGTGTGCTACTTACAGACAGAGGCGCATATTTCACTTATCGAAATGTGAGCAATC 1481

QY 1507 GTCTCTGATACAGAAAGACGAGCGTGTCTGAGAGAAAGTGGGTTAATGTTCAAT 1566
 Db 1482 GTCTCTGATACAGAAAGACGAGCGTGTCTGAGAGAAAGTGGGTTAATGTTCAAT 1541

QY 1567 GATCTGCAAGAGAGGCTCAAGGTTAATGATGCTGCTGATGTCATGATAGCGATGCA 1626
 Db 1542 GATCTGCAAGAGAGGCTCAAGGTTAATGATGCTGCTGATGTCATGATAGCGATGCA 1601

QY 1627 GGCACCTGTAAGAGCCTTTCATCAGCGCTTCTTGAAGACGATGCTTATCAAGTCA 1686
 Db 1602 GGCACCTGTAAGAGCCTTTCATCAGCGCTTCTTGAAGACGATGCTTATCAAGTCA 1661

QY 1687 AGCGAGAACCTATCGTGGGTGATTCACAGTAGGCTTGTGGGTCCAGCGCTGATC 1746
 Db 1662 AGCGAGAACCTATCGTGGGTGATTCACAGTAGGCTTGTGGGTCCAGCGCTGATC 1721

QY 1747 TGTATAGATTCCTCACTGCTCATGAACGCAATGGGCAATGATCATGATCATGTTT 1806
 Db 1722 TGTATAGATTCCTCACTGCTCATGAACGCAATGGGCAATGATCATGATCATGTTT 1781

QY 1807 TGTGTAACAAGATATGTTTTCAGGTTCTCAGCCATGATCTGAAAACTGTGTTCTGT 1866
 Db 1782 TGTGTAACAAGATATGTTTTCAGGTTCTCAGCCATGATCTGAAAACTGTGTTCTGT 1841

QY 1867 TTTAGAACTGTTTGAAGACACAGAGAGCTGAGACACGATTTGCAACAAGATGCAAG 1926
 Db 1842 TTTAGAACTGTTTGAAGACACAGAGAGCTGAGACACGATTTGCAACAAGATGCAAG 1901

QY 1927 CCTGTGATTAATTCOAATGCTGCTTAATCAATATATTAATAACATTAATC 1978
 Db 1902 CCTGTGATTAATTCOAATGCTGCTTAATCAATATATTAATAACATTAATC 1953

RESULT 3
 US-10-425-114-5934
 ; Sequence 5934, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua

```

/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 5934
/ LENGTH: 1533
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: 700553382_FLI
US-10-425-114-5934

Query Match      71.4%; Score 1416.6; DB 12; Length 1533;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 9; Indels 27; Gaps 6;

QY 449 GTTATCCATGTTTCTGAAATCGAAGATGGAATATGTCAAAAGCTTACATATCAGACG 508
DB 1 GTTATCCATGTTTCTGAAATCGAAGATGGAATATGTCAAAAGCTTACATATCAGACG 60
QY 509 GTGATGAATCTGGAATCCAGNATCTGTATACAG----CCTTATGAACTGAGCA 564
DB 61 GTGATGAATCTGGAATCCAGNATCTGTATACAG----CCTTATGAACTGAGCA 119
QY 565 ACTATTGGAAGGTATCGCTATGATGAAAAGCTGACGCTGAGCAAGATGCTACCTGT 624
DB 120 ACTATTGGAAGGTATCGCTATGATGAAAAGCTGACGCTGAGCAAGATGCTACCTGT 179
QY 625 TTCACTTTGGAAGATGATGTCACGAGATTTTTCCTGCTTATTGAAACAAATTCGTGT 684
DB 180 TTCACTTTGGAAGATGATGTCACGAGATTTTTCCTGCTTATTGAAACAAATTCGTGT 239
QY 685 CAAGGACGCGAGTATGACGATTTGATTTGTTTCACTTAACTGATGATTTGTTAA 744
DB 240 CAAGGACGCGAGTATGACGATTTGATTTGTTTCACTTAACTGATGATTTGTTAA 239
QY 745 TCGGATATCTTGAAGCAACTATTCCTGCTTTCGCGAAGAGACTTCTGAGGAGCTGGAT 804
DB 300 TCGGATATCTTGAAGCAACTATTCCTGCTTTCGCGAAGAGACTTCTGAGGAGCTGGAT 359
QY 805 ACAGGATCCAGGATACCTGCTGTTTACTGAGGTTCTTGGGAAAGGCTGGAATCTGGTGC 864
DB 360 ACAGGATCCAGGATACCTGCTGTTTACTGAGGTTCTTGGGAAAGGCTGGAATCTGGTGC 419
QY 865 TGTACTACTTTAGGCGGAGGTGTAGTGACTTGACTGCTGTAACCAATTTGTTAAAGCCTT 924
DB 420 TGTACTACTTTAGGCGGAGGTGTAGTGACTTGACTGCTGTAACCAATTTGTTAAAGCCTT 479
QY 925 GGGACTGAGAGAAATCAGGTATGGAAGAAGTGTGATGTTGTTACTTACTTGTATCCAAA 984
DB 480 GGGACTGAGAGAAATCAGGTATGGAAGAAGTGTGATGTTGTTACTTACTTGTATCCAAA 539
QY 985 TATCTAACCAATGCAAGAGCTGTTCCATCTTAACTTTGAAGGCGCACAGAACTTGC 1044
DB 540 TATCTAACCAATGCAAGAGCTGTTCCATCTTAACTTTGAAGGCGCACAGAACTTGC 599
QY 1045 TTATTTTGGAGCTCAGGTTTTCATCTGCAATCGATGAGAGCCTGCTAGAGAGGTATAT 1104
DB 600 TTATTTTGGAGCTCAGGTTTTCATCTGCAATCGATGAGAGCCTGCTAGAGAGGTATAT 659
QY 1105 TCCAGTTAGGTTAAGATTCATACACCTTAAAGCTCCAGGACCTTATATCCAGACA 1164
DB 660 TCCAGTTAGGTTAAGATTCATACACCTTAAAGCTCCAGGACCTTATATCCAGACA 719
QY 1165 AAGAGCATGATGAAGGTCTGTGTATCTACTAGTGAATAGGCTCAAGTCAAAATGTGAC 1224
DB 720 AAGAGCATGATGAAGGTCTGTGTATCTACTAGTGAATAGGCTCAAGTCAAAATGTGAC 773

```

```

QY 1225 TATGTGACATGTGAGCACTGCGATGCTTGTGCTAGTATGTTTCTGCGAAGGATATC 1284
DB 774 TATGTGACATGTGAGCACTGCGATGCTTGTGCTAGTATGTTTCTGCGAAGGATATC 833
QY 1285 AGGTATTTGCTATTTGAAGATCAGTATATCTGCGAGATGTTGTTGCTACAGCAAGT 1344
DB 834 TGTATAT-----TTGAAGATCTATGATATCTGAGATTTGTTGCTACAGCAAGT 887
QY 1345 TAGTGTTCCTGTGCTGCTGATGATGCAAGATCTGAGATGAGAACTGATACAGAGGC 1404
DB 888 TAGTGTTCCTGTGCTGCTGATGATGCAAGATCTGAGATGAGAACTGATACAGAGGC 945
QY 1405 AAGTGAATCTGACATGATGTTGAAGAGCTTGAAGAAATGCAATGTTGCTTACTTCA 1464
DB 946 ---GAAGTTACCATGATGATGTTGAAGAGCTTGAAGAAATGCAATGTTGCTTACTTCA 1001
QY 1465 GCAGAGGCGGATTAATTTCACTTATCGGAAATGAGCAATCGTCTGTATACAGAAA 1524
DB 1002 GCAGAGGCGGATTAATTTCACTTATCGGAAATGAGCAATCGTCTGTATACAGAAA 1061
QY 1525 GACGGAGCTGTGCTGAGGAAAGTGGGTTTATGTTCAATGATCTTCCAGAGAGCTGC 1584
DB 1062 GACGGAGCTGTGCTGAGGAAAGTGGGTTTATGTTCAATGATCTTCCAGAGAGAGCTGC 1121
QY 1585 AAAGGTTACATGTCGCTGATGATGTCATGATGAGATGCAAGAGCACTGTAAGAGCCCT 1644
DB 1122 AAAGGTTACATGTCGCTGATGATGTCATGATGAGATGCAAGAGCACTGTAAGAGCCCT 1181
QY 1645 TCATGAGCGCTTCTTTGAAGCGATGCTCATACAAATGCAAGCGGAGACCTTATCGT 1704
DB 1182 TCATGAGCGCTTCTTTGAAGCGATGCTCATACAAATGCAAGCGGAGACCTTATCGT 1241
QY 1705 GGGCTGATCAACGATGAGCTTGTGCTGGTCCAGGCGGTATCTGTTATGATTTCCACTC 1764
DB 1242 GGGCTGATCAACGATGAGCTTGTGCTGGTCCAGGCGGTATCTGTTATGATTTCCACTC 1301
QY 1765 GCTTCATGAACGCGATGAGGATGATCATTTGATCATGTTTCTGTTGAACGTTGATGAT 1824
DB 1302 GCTTCATGAACGCGATGAGGATGATCATTTGATCATGTTTCTGTTGAACGTTGATGAT 1361
QY 1825 CTTCAGGTTCTGAGCAAGTATGATGCAAACTGTTTCTGTTGAACGTTGATGAT 1884
DB 1362 CTTCAGGTTCTGAGCAAGTATGATGCAAACTGTTTCTGTTGAACGTTGATGAT 1421
QY 1885 CACCA-----GTAGCTGAGACCGATGATGTCACCAAGATGCAAGCTGTGATATAT 1940
DB 1422 CACCAAGCTGAGCTGAGACCGATGATGTCACCAAGATGCAAGCTGTGATATAT 1481
QY 1941 CCAAGCTGCTATCAATATATATATATTAACAATTATCAATATCT 1985
DB 1482 CCAAGCTGCTATCAATATATATATATTAACAATTATCAATATCT 1526

RESULT 4
US-10-425-114-33565
/ Sequence 33565, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 33565
/ LENGTH: 1115

```


TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: Clone ID: UC-ZMFLMO17134C03_FLI
 US-10-425-114-33565

Query Match 41.9%; Score 831.4; DB 12; Length 1115;
 Best Local Similarity 97.0%; Pred No. 3.9e-250;
 Matches 897; Conservative 0; Mismatches 6; Indels 22; Gaps 4;

1058 CAGGTTTGCATCCACATCGATGAGACCTGCTAGAGAGAGTATATTCAGTTAGGGTT 1117
 209 CAGGTTTGCATCCACATCGATGAGACCTGCTAGAGAGAGTATATTCAGTTAGGGTT 268
 1118 AAGATTTCATACACCTTAAGCTCCAGGCACTTATTAACAGCAAGAGACATGGAT 1177
 269 AAGATTTCATACACCTTAAGCTCCAGGCACTTATTAACAGCAAGAGACATGGAT 328
 1178 AAGGCTGCTGTTTACTACTAGCATAGTCTCAAGTCAATGTCATATGTTGGACAT 1237
 329 AA-----GGTTTACTACTAGCATAGTCTCAAGTCAATGTCATATGTTGGACAT 382
 1238 GTGAGCACTCGGATGCTTGGTCAATGTTTCTGGCAAGGGTATTTGCTATAT----- 437
 383 GTGAGCACTCGGATGCTTGGTCAATGTTTCTGGCAAGGGTATTTGCTATAT----- 437
 1298 ATTGAAGATCTATGATCTGATCTGATGATGTTGTTGCTACAGAGAGAGTATGTTCTGTG 1357
 438 -TTGAAGATCTATGATCTGATGATGTTGTTGCTACAGAGAGAGTATGTTCTGTG 496
 1358 TCACTTGATTCATCAAGATCTGGAGTAGGAACTGATACAGAGGCAAGTGAATTGAC 1417
 497 TCACTTGATTCATCAAGATCTGGAGTAGGAACTGATACAGAGGCAAGTGAATTGAC 550
 1418 CAGTGATGTAAGAGCTGAGAAATAGCAATGTTCTGCTACTCAGAGAGGCGCATTA 1477
 551 CAGTGATGTAAGAGCTGAGAAATAGCAATGTTCTGCTACTCAGAGAGGCGCATTA 610
 1478 ATTCACTTATCGAAATGTGAGCAATGCTCTGATATCTAGAAAGCGGAGCGTGTG 1537
 611 ATTCACTTATCGAAATGTGAGCAATGCTCTGATATCTAGAAAGCGGAGCGTGTG 670
 1538 CTGAGGAAAGTGGGGTATATGTTTCAGATGATCTGCAAGAGCGGCAAGGTTAATG 1597
 671 CTGAGGAAAGTGGGGTATATGTTTCAGATGATCTGCAAGAGCGGCAAGGTTAATG 730
 1598 TCGCTGATGTCATGATAGCGATGCAAGGCACTGTAAGAGCCCTTCATCAGGCGTTG 1657
 731 TCGCTGATGTCATGATAGCGATGCAAGGCACTGTAAGAGCCCTTCATCAGGCGTTG 790
 1658 TTTGAAGACATGTCCTATACAGAGTGAAGCGGAGAACTCTGCTGGGCGTATCAAG 1717
 791 TTTGAAGACATGTCCTATACAGAGTGAAGCGGAGAACTCTGCTGGGCGTATCAAG 850
 1718 TAGGCTTGTCTGGGTCAGGCGGTATATCTGTTATAGATTCCTCAGCTCGCTCATGAAG 1777
 851 TAGGCTTGTCTGGGTCAGGCGGTATATCTGTTATAGATTCCTCAGCTCGCTCATGAAG 910
 1778 GCATGGGCAATTTGATCATGATGTTTGTGTTGAACAAGATGCTTCCAGAGTTCTC 1837
 911 GCATGGGCAATTTGATCATGATGTTTGTGTTGAACAAGATGCTTCCAGAGTTCTC 970
 1838 AGGCAATGATGCAAACTGTTCTGTTTGAAGCTGTTTGCACAGACCA-----GTGA 1893
 971 AGGCAATGATGCAAACTGTTCTGTTTGAAGCTGTTTGCACAGACCA-----GTGA 1030
 1894 GCTGGAGACCGGATTTGCAACAAGATGCGAAGCTGCTGATATTAATCCAACTGCTCTA 1953
 1031 GCTGGAGACCGGATTTGCAACAAGATGCGAAGCTGCTGATATTAATCCAACTGCTCTA 1090
 1954 ATCAATATATATATATAAATATATC 1978
 1091 ATCAATATATATATAAATATATC 1115

RESULT 5
 US-10-425-114-6176
 Sequence 6176, Application US/10425114
 Publication No. US20040034888A1
 GENERAL INFORMATION:
 APPLICANT: Liu, Jingdong
 APPLICANT: Zhou, Yihua
 APPLICANT: Kovalic, David K.
 APPLICANT: Screen, Steven E.
 APPLICANT: Tabaska, Jack E.
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 FILE REFERENCE: 38-21(5313)B
 CURRENT APPLICATION NUMBER: US/10/425,114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 6176
 LENGTH: 1780
 TYPE: DNA
 ORGANISM: Glycine max
 FEATURE:
 OTHER INFORMATION: Clone ID: 700556108_FLI
 US-10-425-114-6176

Query Match 37.6%; Score 745.4; DB 12; Length 1780;
 Best Local Similarity 72.6%; Pred. No. 6.4e-223;
 Matches 1010; Conservative 0; Mismatches 367; Indels 15; Gaps 3;

284 GTGATGAATTCGGGGGCTCTCGGTGCTGCTGCGCCGAGAGATGCTGAGGTGCGCGG 343
 48 GTGATGAATTCGGGGGCTCTCGGTGCTGCTGCGCCGAGAGATGCTGAGGTGCGCGG 107
 344 CTCATCTGACGTTCCCGAGAGAGCGCCCGTCTGCTCTCTGCTGCTAGTGGGAAACC 403
 108 CTATATTAAGTTTCCGAGAGAGAGGCTATGTTGTTCTCTGCTGCTAGTGGGAAACA 167
 404 ACCACACCTTCTCTCTGCTGAGAGAGAGGAGGTGAGATTAATCACTTTCT 463
 168 ACACACACCTTCTCTCTGCTGAGAGAGAGGAGGTGAGATTAATCACTTTCT 227
 464 GAATGGAAGAGGGAATATGCTCAAGAGCTTACATCAAGAGGTGAGTGAATTGGA 523
 228 AGATTTGAGAGAGCTTCTTATTAAGAGCTGATCAAGAGCTGAGATGAGTCACTTGGT 287
 524 CTTCAGAGATCTGTAATCA--GCCCTTAATGAACTGAGCACTATTTGAAGATATC 580
 288 GTGAGCGGATCTGTAATTTCAAGAGCATGAGAGATTTGAGCACTTCTGAAGGGGATA 347
 581 GCTATGATGAAGAGAGTGAAGCTGAGCAAGTCAAGAGCTATTTAGTCTCTTTGGAATGC 640
 348 GCTATGATGAAGAGATTTGAAGAGCTGAGCAAGTCAAGAGCTATTTAGTCTCTTTGGAATGC 407
 641 ATGTCCACAGAGATTTTCTGCTATTTGAACAAATTCGTGCAAGGCAAGGAGATAT 700
 408 ATGTCCACAGAGATTTTCTGCTATCTATTAATAAATAGGTGTCAGAGCTGCGCAATAT 467
 701 GAGGATTTGATATGTTGCTTCAATCACTGATGAAATTTGGAATGAGGATATCTTGA 760
 468 GATGCAATTTGATGTTGTTTATTAACAATGAGCTTCAACAATGCGGCACTTTTGA 527
 761 GCAACTTATCTGCTGTTTGAAGAGCTTCAAGAGCTGAGATGAGATCAAGATCAAGGATA 820
 528 GCAACTTATCTGCTGTTTGAAGAGCTTCAAGAGCTGAGATGAGATCAAGATCAAGGATA 587
 821 CTTGTTTATCTGAGTTTCTTGGAGAGGCTGGAATGTTGTTGTTGTTGTTGTTGTTG 880
 588 GCAATTTATCTGAGTTTCTTGGAGAGGCTGGAATGTTGTTGTTGTTGTTGTTGTTG 647
 881 CGAGTGTATGATCTGATCTGATCAACATTTGTTAAAGCTTGGGACTGAGAGAAAT 940

```

Db      648 AGAGGGGCGAGTATTGACGGCTACAACAATTGGGAAAGCATAGGGTTGCTGAGATC 707
QY      941 CAGGTATGGAAGATGTTGATGTTGACTTACTTGTGATCCAAATATCTACCCACATGCA 1000
Db      708 CAGGTATGGAAGATGTTGATGTTGACTTACTTGTGATCCAAATATCTACCCACAAAGCA 767
QY      1001 AAGACTGTTCCATCTACTTAACATTTGAAGAGCCACAGAACTTGCTTATTTGGTGTGAG 1060
Db      768 GAACCTGTTCTTATTTGACATTTGATGAGGCTGACAGAACTTACTTGGTGTGAG 827
QY      1061 GTTTGATCCCAATTCGATGAGACCTGCTAGAGAGGTGATATTCAGTTAGGGTTAG 1120
Db      828 GTTTCATCCACAGCTATGAGACCTGCTAGAGAGAGTATATTCCTGTTAGGGTTAA 887
QY      1121 AATTCAATCAACCTTAAAGCTCCAGGACCTTATTTACCAACAAGAGCATGATTAAN 1180
Db      888 AATTCTTACCAACCTTAAAGCTCCAGGATCTCTCATCACCAAGGACAGATATGAGCAA- 946
QY      1181 GGTCTGTTGTTACTACTAGCATAGTCTCAAGTCAATGTCATATGTTGAGATGTTG 1240
Db      947 -----GGCAGATTTAAACAAGCATGTTTGAACCTTAATGACATGTTGATATAGCA 1001
QY      1241 AGCACTCGAGTGTGTTGTCAGTATGTTTCTGCGAAGGGTATCAGTATTTGCTATAT 1300
Db      1002 AGCACTCGAGTGTGTTGTCAGTATGTTTCTGCGAAGGGTATCAGTATTTGCTATAT 1055
QY      1301 GAAGATCTATGATATCTGTGATTTGTTGCTACAGTGAAGTTAGTTTCTGTGCTCA 1360
Db      1056 GAAGAGTTGAGCATACAGTATGTTGTTGAGCTAAAGTGAAGTCAAGTGTCTTACCA 1115
QY      1361 CTGTATCCATCAAAAGTCTGAGTAGAGGAGTATGATACAGACAGGCAAGTGAATGACAT 1420
Db      1116 CTGTATCCATCAAAAGTATGAGAGAGAGAGTATTCAGACAGGCAAGTGAATGACAT 1175
QY      1421 GTATGTTGAAGAGTTGAGAAATAGCAATTTGTTGCTTACTTACAGAGAGGCGATAT 1480
Db      1176 GTTGTAGAAAGAACTCAGAAATTCGCTGTGTAATCTCTCAAGAAATGATTCATATC 1235
QY      1481 TCACCTATCGAAATGTGAGCAATCGTCTGATACTAGAAAGAGCGGAGCTGTGCTG 1540
Db      1236 TCTCTCATGGAATTTTCAAGATCATCATATATTTGGAAGAGCTTTCGCTGCTT 1295
QY      1541 AGGAAAGAGTGGGTTATGTTCAATGATCTGCAAGAGGCTCAAGAGTTAAATGTCG 1600
Db      1296 CGAACCTTGGGCTCACTGTGCAATGATCTCCAGGGTGCACTTAAAGTAAACATCTCA 1355
QY      1601 CTGATATGCATGATAGCATGCAAAAGGCACTGTAAGAAAGCCCTTCAAGAGGCTTCTT 1660
Db      1356 TTGTTGTTAAATGACGTGAACAGACAGCATGTGTGAGAGGCTTCTCACTCAGCTTCTT 1415
QY      1661 GAAGACGATGTC 1672
Db      1416 GAGAGTGAAGCTC 1427

```

RESULT 6

```

US-10-424-599-28065
; Sequence 28065, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 28065
; LENGTH: 2882
; TYPE: DNA

```

```

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MR1847_125343C.1
US-10-424-599-28065

Query Match      37.3%; Score 740.6; DB 12; Length 2882;
Best Local Similarity 72.3%; Pred. No. 3e-221; Indels 15; Gaps 3;
Matches 1007; Conservative 0; Mismatches 370

QY      284 GTGATGAAGTTCGGGGGGTCTCGTGTGTCGCGCCGAGAGATGAGTGGTGGCCG 343
Db      918 GTCATGAAGTTTGGGGGCTCTGTGCTTCTGCTGATAGATGAAGAGGTGGCTACC 977
QY      344 CTCATCTGACGTTCCCGAGAGAGCCCGCTGTGCTTCTCTCCATCGGGGAAAC 403
Db      978 CTTATATTGAAGTTTCCCGAGAGAGGCTTATGTTGTTCTCTCTGCTATGGGAAACCA 1037
QY      404 ACCAACCACTTCTCTGCTGTGAGAGAGGCAAGTATGAGTTATCCATGTTTCT 463
Db      1038 ACAACCAAGCTTTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1097
QY      464 GAATTCGAAGTGAATATGTCGAAAGCCTTACATATCAAGCGGTGATGAATCTTGA 523
Db      1098 AGTATTTGAGAGAGCTTTGCTTATTAAGACCTGCTATCTAGAGCTGTGATCAGTTGCT 1157
QY      524 CTTCAGNATCTGTAATACAA---GCCTTATGAACTGAGAGCACTATTGAAGGTATC 580
Db      1158 GTGAGCGATCTGTTATTTCAAGCATCTGAAGAAATGAGAGCACTTCGAAGGGGATA 1217
QY      581 GCTATGATGAAGAGCTGACGCGCTAGAGCAAGTATGATCCTGTTTCATTTGAGATGC 640
Db      1218 GCTATGATGAAGAGATGATGACTTAAAGAGCTCAGACATATTAATCTCTTGGAGATGC 1277
QY      641 ATGTCCACAGAGATTTTTCGCTTATTTGAACAAATTTGCTGCAAGGACGCGAGTAT 700
Db      1278 ATGTCCACAGAGATTTTTCGCTTATTTGAACAAATTTGCTGCAAGGACGCGAGTAT 1337
QY      701 GACGCAATTTGATTTGTTGTTTCAATTAACATGATGAATTTGTTATGCGATATCTTGA 760
Db      1338 GATGCAATTTGATTTGTTTCAATTAACATGATGAATTTGTTATGCGATATCTTGA 1397
QY      761 GCAACCTATCCGCTGCTGGAAGAGACTCATAGGGAGCGATACAGATCCAGAGATA 820
Db      1398 GCAACCTATCCGCTGCTGGAAGAGACTCATAGGGAGCGATACAGATCCAGAGATA 1457
QY      821 CCTGTTGTTACTGAGTTCTCTGGAAGAGGCTGGAATCTGCTGCTGATCTTAAAGG 880
Db      1458 GCAATTTGTTACAGGCTTCTTGAAGAGCCCGGAAATCATGTGACAGTCAACATGGGT 1517
QY      881 CGAGGTGTGATGACTTGCTGCTACCAACATTTGTTAAACCTTGGGACTGAGAGAAAT 940
Db      1518 AGAGGGGCGAGTGAATGACGCTAACCAATTTGGAAGACCTAGAGTTGCTGAGATC 1577
QY      941 CAGGTATGGAAGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATG 1000
Db      1578 CAGGTATGGAAGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATG 1637
QY      1001 AAGACTGTTCCATCTACTTAACATTTGAAGAGCCACAGAACTTGCTTATTTGGTGTGAG 1060
Db      1638 GAACCTGTTCTTATTTGACATTTGATGAGGCTGCAAACTTACGTAATCTTGGTGTGAG 1697
QY      1061 GTTTGATCCACATGATGATGAGACCTGCTAGAGAGGTATTCAGTTAGGGTTAG 1120
Db      1698 GTTTCATCCACAGCTCTATGAGACCTGCTAGAGAGGTATTCAGTTAGGGTTAA 1757
QY      1121 AATTATACAAACCTTAAAGCTCCAGGACCTTATTTACCAACAAGAGCATGATTAAN 1180
Db      1758 AATTCTTACAAACCTTAAAGCTCCAGGATCTCTCATCCAGGACAGATATGAGAACAA- 1816
QY      1181 GGTCTGTTGTTACTACTAGCATAGTCTCAAGTCAATGTCATATGTTGAGATGTTG 1240
Db      1817 -----GGCAGATTTAAACAAGCATGTTTAAAGATATGATGACATGTTGATATAGCA 1871

```

QY	1241	AGCACTCGGAAGCTTGGTCAGTATGGTTTCTGGCAAGGATACAGGATTTGCTATAT	1300
Db	1872	AGCACTCGCAATGCTTGGTCAGTATGGTTTCTGGCAAGGATACAGGATTTGCTATAT	1925
QY	1301	GAAGATCTATGATATCTGTGGATTGGTTGCTACAGCTGAAGTTAGTCTGTGTCA	1360
Db	1926	GAAGAGTTTAGGCATATCAGTTGATTTGTATCTCAAGTGAAGCAATGTTTCTGTACA	1985
QY	1361	CTTATGTCATCAAAAGATCTGGAGTGAAGGAATCTGATACAGCAGGCAATGAACTTGAACAT	1420
Db	1986	CTGATATCCATCAAAAGCTATGGAGCAGAGAGCTATTTACAGCAGCAATGAACTTGAACAT	2045
QY	1421	GTAATTGAAGAGCTTGGAGAAATAGCAATTTGTTGTTACTTACAGCAGAGGCGATATAT	1480
Db	2046	GTTGTATGAAGCACTCGAGAAAATCGCTGTGTGATCTCTACAGAAATGATCATATC	2105
QY	1481	TCACTTATCGGAAATGTGAGCAATCGTCTCTGATACAGAAAGACGGACGTGTCTG	1540
Db	2106	TCTCTCATTTGAAATGTTTCAAGATCATCATATATTTGGAAGAGGCTTCCGTTCTT	2165
QY	1541	AGGAAAGATGGGGTTATATGTTCAATGATCTCCGAAAGAGCGTCAAAAGTTAATCTCG	1600
Db	2166	CGAACCCCTGGCGTCACTGACGTCAATGATATCTCCAGGGTGATTTAAGTGAATCTCA	2225
QY	1601	CTGATATGTCATGATAGCATGCAAGGCACTCGTAGAGCCCTTCATCAAGGCTTCTT	1660
Db	2226	TTGTTTGAATGACAGTGAAGCAGAGCAGTGTGAAGGCTCTCCACTAGCTTCTT	2285
QY	1661	GAAGACGATGTC	1672
Db	2286	GAGAGTGAGCTC	2297
RESULT 7			
US-09-938-842A-1744			
Sequence 1744, Application US/09938842A			
Patent No. US20020160378A1			
GENERAL INFORMATION:			
APPLICANT: Harper, Jeff			
APPLICANT: Kieps, Joel			
APPLICANT: Kieps, Xun			
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING			
FILE REFERENCE: SCRIPI300-3			
CURRENT APPLICATION NUMBER: US/09/938,842A			
PRIOR FILING DATE: 2001-08-24			
PRIOR APPLICATION NUMBER: US 60/227,866			
PRIOR FILING DATE: 2000-08-24			
PRIOR APPLICATION NUMBER: US 60/264,647			
PRIOR FILING DATE: 2001-01-16			
PRIOR APPLICATION NUMBER: US 60/300,111			
PRIOR FILING DATE: 2001-06-22			
NUMBER OF SEQ ID NOS: 5379			
SEQ ID NO 1744			
LENGTH: 1710			
TYPE: DNA			
ORGANISM: Arabidopsis thaliana			
US-09-938-842A-1744			

Query Match:	35.6%	Score 707.6	DB 9	Length 1110
Best Local Similarity	70.5%	Pred. NO. 4,9e-211		
Matches 1008	Conservative	0	Mismatches 401	Indels 21
			Gaps	4

QY	254	ATGAGGAGATTGGGGGATCAGCTCAGCGGTGATGAATTCGGGGGCTCTCGGTGCG	313
Db	235	ACGAGGTAGATTAGAAAGGTATCACTGTCGTGATGAAGTTGGTGGATTTTCGGTGGG	294
QY	314	TGGGCGCGAGGATGGCTGAGGTGGCGGCTCATCTGACGTTCCCGAGAGCGCCCC	373
Db	295	TCAGCTGAGGATGAAAGGAAGTGGCTGATTTGATTTGACTTTCCGGAAGAAATGCC	354
QY	374	GTCGTGCTTCTCTTCGCAATGGGAAAACCAACCAACCTTCTCTTGTGAGAGAAAG	433

Db	355	GTCTATTGTTCTCTCTGCTATGGGAAACTTACCAACATCTCTTGCTTGGGAGAGAG	414
Qy	434	GCAATBAGGTGTGAGTTATTCATGTTTCTGAAATGGAAAGTGGAAATBGTCAAAGC	493
Db	415	GGGGTTAGTGTGGTGTGTTCTTAAGCATCGAATGAGAGATGAGCTTATAAGGAA	474
Qy	494	CTACATATCAAGACGGTGTGATGAACTTGGACT---TCCAAAGATCTGTAAATCAAGCCTT	550
Db	475	TTCGATATTCAGGACCGGTAAAGAGCTCAACATGATCCCTCTGTTATTTGACCTATTGG	534
Qy	551	TATGAACTGAGCAACTATTGAAAGGTATCGCTATGATGAAAGAGCTGACGCTGAGCC	610
Db	535	GAGGAATCTGAGCAACTCTGAAAGGCAATGGCCATGATGAAGAAATTGACATTCGAACC	594
Qy	611	AGTGACTACCTGTTTCAATTGGAGAAATGATGTCCAGCAGAGATTTTCTCTTATTTG	670
Db	595	AGAGATTACTTATGTCCTTTTGGAGATGTTTGTCTACAGAAATTTTGCTCTTATCTT	654
Qy	671	AACAAAATTCGTGTCAAGGACGCGAGATGACGATTTGATTTGTTTCATTACAAC	730
Db	655	AATACATCGAGTGTCAAGGACGCGCAATATGATGACATTTGAATTTGTTCTTACAACG	714
Qy	731	GATGATTTTGGTATGCGGATATTTTGAAGCAACTATCCGTCTGTGCGAAGAGACTT	790
Db	715	GATGATTTTCAAAATGGGATATCTGGAAGCAACTTATCCAGCTGTGTGCAAGGATTA	774
Qy	791	CATGGGACCTGGATPCAGGATCCAGCATACCTGTTGTTACTCGGTTCTTGGAGAGGC	850
Db	775	TATGATGATGATGATGATGATGATCCGCTGTTCTGATTTGTAACAGTTTCTTGGAAAGGT	834
Qy	851	TGGAATCTGTGCTCTTACTTATTTAGCCGAGGTGTATGTGACTTGCTTACAAAC	910
Db	835	TGGAAACCTGTGCGGTACTACCTTGGAGGGGTGGAGATTTGACGGCAACCA	894
Qy	911	ATTGGTAAAGCTTGGAGCTGAGAAATTCAGGTATGAAAGATTTGATGTTGTTACTT	970
Db	895	ATTGGTAAAGCTTGGGTGTTGAAGAGATTCAGGTGTGAAAGATGTGATGTGTTCTTA	954
Qy	971	ACTTGATCCAAATATCTACCCATGCAAGACTGTTCCATCTTAACTTTGAAGAG	1030
Db	955	ACATGTGACCCACTATATATAACGAGCTACACAGTACCATATCTGACATTTGACGAA	1014
Qy	1031	GCCACGAACCTTGCTTATTTTGGTGCAGGTTTGCATCCCAATCCATGAGTGAACCTGCT	1090
Db	1015	GACGCGAGCTAGCTTATTTTGGTGCAGAGCTTTCGACCCAGTCAATGAGACGACGA	1074
Qy	1091	AGAGAAGTGAATATCCAGTTAGGTTAAGATTCATACAACTTAAAGCTTCAGGACCC	1150
Db	1075	AGAGAGGGTGAATTCCTGTTAGGTTAAAAATTTCTTATACCTTAAAGGCTCTGGAAC	1134
Qy	1151	CTTATTTACAGCAAAAGAGACATBGTGAATBANGGTCTGTTGTACTAATGACATAGTGTCT	1210
Db	1135	ATCATCTCAATAAACAAGACATGACCAAG-----TATTTCAACGAGATTTGTTGTG	1188
Qy	1211	AAGTCAATATGCTACATATTTGGAATGTGAGCACTCGATGCTTGGTACGATAGTGT	1270
Db	1189	AAACGTAATGTACCATGCTGGATATATGACAGCAACCGAATGTTGGTCAATGTGGCTT	1248
Qy	1271	CTGGCAGGGATTCAGTATTTGCTATTTGAAAGATCTATGTATCTGTGATTTGTGTT	1330
Db	1249	CTTGCAGAAAGTATTCGATAT-----TTGAAGACTTGGCAATTCCTGATATGTTGTT	1302
Qy	1331	GCTACAGAGATGATGATGTTCTGTGTCACTTGATCCATCAAAGATCTGAGTGGGAA	1390
Db	1303	GCCACTAGTGAAGTATGATATCTCTGACACTGATCTTCAAAATTTTGGAGCGAGAA	1362
Qy	1391	CTGATATCAGACGACGAAGTGAACCTTGACCATGTATGTAAGAGCTTGAGAAATTAACAAT	1450
Db	1363	CTGATTTCAACAG-----GAGCTTGATCAATGTATGTAAGAACTGAGAAATTTGACGTT	1416
Qy	1451	GTTTCGTACTTCTACACAGAGGCGATTAATTTCACTTATCGGAAATGTGAGCAATCGCTCT	1510

Db 1417 GTGAATCTTCTTAAAGAGAGCAATCTCTCTAATTGGAAATGTTCAACATTCCTCC 1476
 QY 1511 CTGATACCTAGAAAGACGGGACGTGTGAGAGAAAGTGGGTTAATGTTACAGATATC 1570
 Db 1477 CTGATTTTAGAGAGGCGGTTTCATGTTCTTTATACAAAGGTGCATATGTCAGATATA 1536
 QY 1571 TCGCAAGAGCGCTCAAGGTTAATCATGTGCTGATAGTCCATGATAGCAAGGCA 1630
 Db 1537 TCACAAGAGCATCCAGGTAACATTTCTTTATAGTAAACGAAGCTGAAGCTGAAGGT 1596
 QY 1631 CTCGTAGAGCCCTTCATCAGGCGTTCTTTGAAGCATGTCTCATCA 1680
 Db 1597 TGTGTTCAAGCTCTTCAAAATCCTTCTTCAGAGCGGTGACCTCTCAGA 1646

RESULT 8

US-09-938-842A-1744
 ; Sequence 1744, Application US/0993842A
 ; Publication No. US20040009476A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kieps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1744
 ; LENGTH: 1710
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-1744

Query Match 35.6%; Score 707.6; DB 11; Length 1710;
 Best Local Similarity 70.5%; Pred. No. 4,9e-21;
 Matches 1008; Conservative 0; Mismatches 401; Indels 21; Gaps 4;

QY 254 ATGAGAGATGGGGATCAGCTCAGCGTGTGATGAGTTGCGGGGTCCTCGATGCG 313
 Db 235 ACGGAGGTAGTGAAGAGGTATCACTGCGTGAAGATTGGTGGATCTTGGGTGGC 294
 QY 314 TCGGCCGCGAGATGAGTGGTGGCGGCTCATCTGACGTTCCCGAGAGCGCCC 373
 Db 295 TCAGCTGAGAGATGAAGAGTGTCTATTTGATTTTGACTTTTCGGAAGAAAGCC 354
 QY 374 GTGCGTCTCTCTGCGCATGGGAAAAACCAACAACACTTCTCTGCTGAGAGAG 433
 Db 355 GTCATTTGTTCTCTCTGCTATGGGAAAACTACCAACAATCTCTTGCTGGCGGAGAG 414
 QY 434 GCAGTAGGCTGTGAGTATTCATGTTTCTGAATCGAAGATGGAATATGCTCAAAAGC 493
 Db 415 GCGGTAGTTGGTGGTCTTCTAATGATCTGAGATTGAGAGTTCATATAAGAA 474
 QY 494 CATACATCAAGCGGTGATGACTGGAATCTTCAAGNATGTAAATACAGGCTT 550
 Db 475 TTGCATATCAGACGTGAAGAGCTCAACATGATTCCTCTGTTATTTTGAACCTATTG 534
 QY 551 TATGAACCTGAGCACTATTGAAGGTATCGCTATGATGAAGAAAGCTGACGCTAGAAC 610
 Db 535 GAGGAACCTGAGAACTCTCGAAAGCATTCGCAATGAGAAATGACACTTCGAAAC 594
 QY 611 AGTGACTACCTGTTTCAATTTGAGAGATGATGTCACACAGATTTTTCGCTATTTG 670
 Db 595 AGAGATTACTGTACTCTTTTGAAGAGTGTGTTGTCACAGGATTTTTCGCTTATCTT 654

QY 671 AACAAATTCGTGTCAAGGACGCGAGATATGACGATTTGATNTTGGTTTCAATTAAC 730
 Db 655 AATCAATCGGTGTCAAGGACGCGCAATATGATGCAATTTGMAATTTGTTTCAATTAAC 714
 QY 731 GATGAATTTGGTAAATGGGGAATATCTTGAAGCAACCATCTGCTGTGGAAGACTT 790
 Db 715 GATGATTTCACAAAATGGGGAATATCTGAGGCAACTTATTCAGCTGTTGCCAAGATTA 774
 QY 791 CATGGGACTGATACAGATTCAGGCAATCTGTTGTTAATCGGTTCTCTGGAAAGGC 850
 Db 775 TATGATGATGATGATGATGATCTGCTGTTCTTATTTGAACAGGTTCTCTGGGAAGGT 834
 QY 851 TGAATAATCGGTGTGTAATCTTATAGCGGAGGTGTGATGATGATGATGATGATGATG 910
 Db 835 TGAATAATCGGTGTGTAATCTTATAGCGGAGGTGTGATGATGATGATGATGATGATG 894
 QY 911 ATTGTAAAGCCTTGGAGCTGAGAAATCAGTATGGAAGATGATGATGATGATGATG 970
 Db 895 ATTGTAAAGCCTTGGAGCTGAGAAATCAGTATGGAAGATGATGATGATGATGATG 954
 QY 971 ACTGTGATCCAAATATCTACCCACATGCAAAAGCTGTTCCATTAATTTGAGAG 1030
 Db 955 AATGTGACCTCTATATATATTAAGCGAGCTACACCGATACATATCTGACATTTGCGAA 1014
 QY 1031 GCCACAGACTGTGCTAATTTTGGTGTGCTCAGGTTTGGATCCACATGATGATGATGATG 1090
 Db 1015 GCAAGCGAGCTAGCTTATTTTGGTGTGCTCAGGTTTGGATCCACATGATGATGATG 1074
 QY 1091 AGAGAGGTGATATTCAGTTAGGTTAAATTCATACCAACCTTAAGCTCCAGGAC 1150
 Db 1075 AGAGAGGTGATATTCAGTTAGGTTAAATTCATACCAACCTTAAGCTCCAGGAC 1134
 QY 1151 CTTATTACCAAGCAAGAGCATGATGATGATGATGATGATGATGATGATGATGATG 1210
 Db 1135 ATCATCACTAAACAAAGACATGACCAAG-----TATTTAAGAGCATGTTGCG 1188
 QY 1211 AAGTCAATGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1270
 Db 1189 AAACGTATGTGACCAAGCTGATGATGATGATGATGATGATGATGATGATGATG 1248
 QY 1271 CTGCAAGGTATCAGGATTTGCTATATTAAGATCTATGATGATGATGATGATGATG 1330
 Db 1249 CTGCAAGGTATCAGGATTTGCTATATTAAGATCTATGATGATGATGATGATGATG 1302
 QY 1331 GCTACAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1390
 Db 1303 GCTACAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1362
 QY 1391 CTGATAGCAGGAGCAAGTGAACCTGACATGATGATGATGATGATGATGATGATGATG 1450
 Db 1363 CTGATAGCAGGAGCAAGTGAACCTGACATGATGATGATGATGATGATGATGATGATG 1416
 QY 1451 GTTGTCTTACTTCAAGAGGCGATTAATTTCACTTAATGGAATGATGATGATGATG 1510
 Db 1417 GTTGTCTTACTTCAAGAGGCGATTAATTTCACTTAATGGAATGATGATGATGATG 1476
 QY 1511 CTGATAGCAGGAGCAAGTGAACCTGACATGATGATGATGATGATGATGATGATGATG 1570
 Db 1477 CTGATTTAGAGAGGCGTTTCAATGTTCTTATATCCAAAGGTGCAATGTCAGATTA 1536
 QY 1571 TCGAAGAGCGGTCAAGGTTAATCATGTGCTGATGATGATGATGATGATGATGATG 1630
 Db 1537 TCACAAGAGCATCCAGGTAACATTTCTTTATAGTAAACGAAGCTGAAGCTGAAGGT 1596
 QY 1631 CTCGTAGAGCCCTTCATCAGGCGTTCTTTGAAGCATGCTTATCA 1680
 Db 1597 TGTGTTAGGCTCTTCAAAATCCTTCTTCAGAGCGGTGACCTCTCAGA 1646

RESULT 9

US-09-890-813-15
 ; Sequence 15, Application US/09890813

QY 573 AAGATATCGCTATGATGAAAGAGCTGACGCCCTAGGACCAAGTACTACTGTTTCATTTG 632
 Db 122 AAGAGATTGCTATGATGAAAGAGCTGACCCCTTAGGACCAAGATTAACCTGTTCCCTTTG 181
 QY 633 GAGAAATGATGCCACCGAGATTTTTCGCTTATTTGACAAATTCGTGTCAAGGAC 692
 Db 182 GTGAATGATATGTCACAAAGATTTTTCGCTGATTTTGAACAACTTGGAAAAAGCAC 241
 QY 693 GGCAGTATGACGCAATTTGATTTGTTTCATTACCACTGATGAATTTGGTAATCGGATA 752
 Db 242 GGCAGTATGATGATTTGATTTGCTTTATTAACCAACGATGATTTTCAAAATCGGATA 301
 QY 753 TCTTAAGAACCACTTCTGCTGTTGGAGAGACCTCATGGGAGCTGATAGATC 812
 Db 302 TCTTAAGCCACTTACCTGCTGTTGGAGAGACCTCATGAACTGATGATGATC 361
 QY 813 CAGCGATACCTGTTTACTGCTGCTTCTTGGAGAGGCTGAAATCTGCTGTAATCTA 872
 Db 362 CTGCTATCTTATGATCACTGCTTCTTGGAGAGGATGTAATCATGCTGCTGACCA 421
 QY 873 CTTTAAAGCCGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 932
 Db 422 CTTTAAAGCCGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 481
 QY 933 GAGAAATTCAGGATGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 992
 Db 482 GAGAAATTCAGGATGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 541
 QY 993 CACATCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1052
 Db 542 CAATATCTATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
 QY 1053 GTGCTAGGTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1112
 Db 602 GTGCTAGGTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
 QY 1113 GGGTAAAGATTCATGAAACCTTAAGCTCCAGGACCCCTTATTCAGACAAAGACAA 1172
 Db 662 GAGTAAAGATTCATGAAACCTTAAGCTCCAGGACCCCTTATTCAGACAAAGACAA 721
 QY 1173 TGATTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1232
 Db 722 TGAGCAAGAG-----CATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 775
 QY 1233 ACATTTGAGCACTCGGATGCTTGTGCTGATGATGATGATGATGATGATGATGATGATGAT 1292
 Db 776 ATATAGTGAAGCAAGAGATGCTTGGCAATACGGGTTTATGCAAAAGATTTCTCATATAT 835
 QY 1293 GCATATTAAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1352
 Db 836 -----TTGAGATTTGGGATCTCTGTTGATGATGATGATGATGATGATGATGATGATGAT 889
 QY 1353 CTGTGTCACTTGAATCCATCAAAAGATCTGAGTAGGAACTGATACAGAGGCAAGTAGAC 1412
 Db 890 CGTTGACTGATGATCAAAAGATCTGAGTAGGAACTGATACAGAGGCAAGTAGAC 943
 QY 1413 TTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1472
 Db 944 TTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1532
 QY 1533 GTGTGTGAGGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1592
 Db 1064 ACCTTTGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1123
 QY 1593 ACATTTGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1652
 Db 1124 ACATTTGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1183

QY 1653 CGTTCTTGAAGACGATGCTTATCAACAGTGAAGGAGAGAACTACTGTTGGCTGAT 1712
 Db 1184 CGTTCTTGAAGACGATGCTTATCAACAGTGAAGGAGAGAACTACTGTTGGCTGAT 1243
 QY 1713 C 1713
 Db 1244 C 1244

RESULT 11
 US-10-425-114-8335
 ; Sequence 8335, Application US/10425114
 ; Publication No. US2004003488A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 8335
 ; LENGTH: 1361
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700747979_FLI
 ; US-10-425-114-8335

Query Match 30.6%; Score 608.2; DB 12; Length 1361;
 Best Local Similarity 70.8%; Pred. No. 8.1e-180;
 Matches 872; Conservative 0; Mismatches 339; Indels 21; Gaps 4;

QY 444 GTGAGATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 503
 Db 1 GTGAGATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
 QY 504 AGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 560
 Db 61 AGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 561 AGCAACTATGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 620
 Db 121 AGCAACTATGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 621 TTGTTCAATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 680
 Db 181 TTGTTCAATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 681 GTGTCAAGCAGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 740
 Db 241 GTGTCAAGCAGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 741 GTATGCGGATATCTTGAAGCAACTATCTGCTGTTGGAGAGACTCATGGGACT 800
 Db 301 GTATGCGGATATCTTGAAGCAACTATCTGCTGTTGGAGAGACTCATGGGACT 360
 QY 801 GGATACAGATCAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 860
 Db 361 GGATACAGATCAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 QY 861 GTGCTGTAATCTTGAAGCAACTATCTGCTGTTGGAGAGACTCATGGGACT 920
 Db 421 GTGCTGTAATCTTGAAGCAACTATCTGCTGTTGGAGAGACTCATGGGACT 480
 QY 921 CCTTGGAGCTGAGAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 980
 Db 481 CCTTGGAGCTGAGAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540

```

RESULT 12
US-09-890-813-1
; Sequence 1, Application US/09890813
; Publication No. US20020183486A1
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: Aspartate Kinase
; FILE REFERENCE: BBI430.PCT
; CURRENT APPLICATION NUMBER: US/09/890,813
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/172944
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 565
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURES:
; NAME/KEY: unsure
; LOCATION: (127)
US-09-890-813-1

```

```

APPLICANT: LA ROSA Jhonatan V
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated wit
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ. ID NOS: 285664
SEQ. ID NO. 28060
LENGTH: 2490
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_125339C.1
US-10-424-599-28060

Query Match      26.9%; Score 533.2; DB 12; Length 2490;
Best Local Similarity 63.9%; Pred.No. 5,1e-156;
Matches 987; Conservative 0; Mismatches 384; Indels 174; Gaps 5;

```


Db 184 AGACCTGTGGAACAGTTGGAGTGGACGAAATGTTATTTGAAAGCATCTTGAAGAAATTG 243
 QY 560 GAGCAACTTATGAAAGATTCGCTATGATGAAAAGCTGAGCGCTTAGAACCAAGACTAC 619
 Db 244 GAGCAACTTATGAAAGGATGATGATGAAAAGGATGACTCCAGGACTCAAGACTAT 303
 QY 620 CTGTGTTCAATTTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 679
 Db 304 TTAGTTCAATTTGAAAGTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 363
 QY 680 CGTGTCAAGGACGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 739
 Db 364 GAGGTTAGGCGCGCGCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 423
 QY 740 GGTATGCGGATATCTTGAAGCAACCTATCTGCTGTGGGAAGAGACTCATGGGAC 799
 Db 424 ACAATGCTGACATTTTGAAGCAACATCTCTGCTGTGGGAAGAGACTCATAGTAT 483
 QY 800 TGGATACAGATTCAGCGATACCTGTTGTTACTGCGTTCTTGGAGGCGCTGGAATCT 859
 Db 484 TGGGTTCTGATCTGCTATTCCTCAATTTGTTACAGGCTTCTTGGGAAGCGAATAATCA 543
 QY 860 GGTGCTGTACTACTTTAGCGCGAGGTGGTAGTGACTGACGCTACCAACCATGGTAAA 919
 Db 544 TGTGCGGTGACACATTTGGGTAGAGTGGAGTGAATTTTAACTGCTACCAACATTTGGGAAA 603
 QY 920 GCCTTGGGACTGAGAAATTCAGGTATGAAAGATGTTGATGTTGTTACTTACTTGTGAT 979
 Db 604 GCGCTAGGTTGCGCAAAATTCAGGTATGAAAGATGTTGATGTTGTTACTTACTTGTGAT 663
 QY 980 CCAAAATCTACCCCAATGCAAAAGCTGTTCCACTTAACTTGAAGAGCCACAGAA 1039
 Db 664 CCAAAATCTACCCCAATGCAAAAGCTGTTCCACTTAACTTGAAGAGCCACAGAA 723
 QY 1040 CTGTCTTATTTTGGTCTCAG----- 1060
 Db 724 CTGTGATCTTGTGCTCAGTGTAAACAAATTTGATAGGAAAGCTCAATACTG 783
 QY 1061 ----- 1060
 Db 784 CAAATGTACACTGGGAAAGAAATAGAAACTGAAAAAGAGCCTACAGGCTAATATTC 843
 QY 1061 -----GTTTGG 1066
 Db 844 AAAAGGCTATTTCTTGACCATGAGAGAAACCGGATAGAAAGTTATCAACAAAGTCTTC 903
 QY 1067 CATCCAAATCGATGAGACCTGCTAGAGAAAGTATTTCCAGTTAGGGTTAAGAAATTC 1126
 Db 904 CATCCAGCTCTATGAGACCGGCTAGGAAAGTATTTCCAGTTAGGGTTAAGAAATTC 963
 QY 1127 TACAACTCTAAAGCTCCAGGACCCCTTATTTACCAAGAAAGAGCATGATGATGATGATG 1186
 Db 964 TACAACTCTAAAGCTCCAGGACCCCTTATTTACCAAGAAAGAGCATGATGATGATGATG 1017
 QY 1187 GTTGTACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1246
 Db 1018 GCATTTATTAACAAGCATTTGTTGAAACGTAATGATGATGATGATGATGATGATGATG 1077
 QY 1247 CGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1306
 Db 1078 CGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1131
 QY 1307 CTATGATATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1366
 Db 1132 TTAGCATATCAAGTTATGTTAGCTACAGTGAAGTATGATTTCTTACATTTGAT 1191
 QY 1367 CCATCAAAAGCTGAGTGGGAAAGTATGATGATGATGATGATGATGATGATGATGATGATG 1426
 Db 1192 CCATCAAAAGCTGAGTGGGAAAGTATGATGATGATGATGATGATGATGATGATGATGATG 1245
 QY 1427 GAAGACCTTGAGAAATGCAATGTTGCTTACTTACAGAGGCGCATTAATTTCACTT 1486

Db 1246 GAAGAACTGGAAGAAATTTGCTGTGTAAATCTCTTAAAGACAGATCATTAATCTCTTC 1305
 QY 1487 ATCGGAAATGTGAGCAATCGCTCTGTATCTAGTAAAGAAAGACGGGACGTGTGAGAGAA 1546
 Db 1306 ATTTGAAATGTTTCAGAAATCATCTAATATTTGAGAAAGGCGTTTCAATGTTCTTCGAACT 1365
 QY 1547 ACTGGGCTTAAATGTTCAATGATCTCGGAAAGAGCTCAAGGTTAAACATGTGCTGATA 1606
 Db 1366 CTGGGATCACTGTTCAATGATCTCTCAGGCTGCACTTAAGTGAATATCTGTTAGTT 1425
 QY 1607 GTCCATGATPAGGATGCAAGAAAGCACTGTAGAAAGCCCTTCATCAGGCGTTCTTTGAAAG 1666
 Db 1426 ATAAATGACATGAAAGACCAAGTGTCTCAGGCTCTCCAAAGGCTTCTTTGAGAGC 1485
 QY 1667 GATGTC 1672
 Db 1486 GAATCTC 1491

RESULT 15
 US-10-425-114-7451
 ; Sequence 7451, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 7451
 ; LENGTH: 1260
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700663367_F11
 US-10-425-114-7451

Query Match 23.7%; Score 470.4; DB 12; Length 1260;
 Best Local Similarity 71.6%; Pred. No. 1.8e-136;
 Matches 666; Conservative 0; Mismatches 246; Indels 18; Gaps 3;

QY 743 AATGCGATATTTTGAAGCAACCTATCTCTGTTGCGAAGAACTTATGAGGACTGG 802
 Db 3 AATGCGCAATTTTGAAGCAACCTATCTCTGTTGCGAAGAACTTATGAGGACTGG 62
 QY 803 ATACAGGATCCAGCATACCTGTTTACTGCGGTTCTTGGGAAGGCTGAAATCTGAT 862
 Db 63 CTCTGATCTCGCAATTCGCAATTTCTTACAGGCTTCTTGGAAAGGCCAGAAATCATGT 122
 QY 863 GCTGTACTACTTTAGCCGAGGTGATGATGATGATGATGATGATGATGATGATGATGATG 922
 Db 123 GAGTGTACAACCTGAGTGAAGGAGGAGTATTTGACAGTACAGCAATTTGGGAAAGCA 182
 QY 923 TTGGGACGAGGAATTCAGTATGAAAGATGTTGATGATGATGATGATGATGATGATGATG 982
 Db 183 CTAGGTTACCTGCAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 242
 QY 983 AATATCTACCATGCAAGAAAGCTGTTCCATTAATTTAAACATTTGAAAGGCCACAGAACTT 1042
 Db 243 AATATATACCCAAAGCCGAAACCTGTTCTTATTTGACATTTGACAGGCTGTGTAATTA 302
 QY 1043 GCTTATTTTGTGCTCAGGTTTGTGATCACAATCGATGAGTGAAGTGTGATGAGTGAAGGAT 1102
 Db 303 GCGTACTTTGCTCAGGTTTGTGATCACAATCGATGAGTGAAGTGTGATGAGTGAAGGAT 362
 QY 1103 ATTCAGTTAGGTTAGAAATTCATACAAACCTTAAGCTCCAGGACCTTATTAACGAGA 1162

Wed Mar 24 14:12:54 2004

us-09-890-813-5.rnpb

Page 14

Db 363 ATTCCGTGATGGTTAAAATTTCTTCAATCTCTAAAGCTCCAGATCTCTATCAACCAAG 422

QY 1163 CAAAGAGCATGGATTAANGTCTGTGTTGTACTAACTTAGCATAGTGTCTCAAGTCAAAATGTC 1222

Db 423 GCAGAGGATATGAGCAA-----GGCAGTGTAAACAAGCAATGTTTTGAAGATATGTG 476

QY 1223 ACTAATGTGGACATGTGAGGACACGAGTGGTGTGTGACATAGTGTCTTGTGCAAGGGTA 1282

Db 477 ACCATGTGGATATGATGATGACTGCGATGCTGTGTGTCATGATAGTGTCTTGTCTAAGGTG 536

QY 1283 TCAGTATTTGCTATATTGAAGATCTATGTATATCTGTGATTTGTGTCTCCACAGTGA 1342

Db 537 T-----TTTCAATCTTTGAAGAGTTAGGACATACAGTTGATGTTGTACTCAAGTGA 590

QY 1343 GTTAGTGTCTTGTGTCTCACTTGATTCATCAAAGTCTGGAGTGGGAAGTGAATACAGAG 1402

Db 591 GTGAGGTGTTCTTTGACATTTGGATTCATCAAAGCTATGGAGCAAGATTAATTCAGAG 650

QY 1403 GCAGTGAAGCTTGACCATGTAGTGAAGAGCTTGAGAAATATAGCAATTTGTGTACTTT 1462

Db 651 -----GACTTGACCATGTTGTGAGAAAGTCTGAGAAATGCTGTGTGATATCTCTG 704

QY 1463 CAGCAGAGGGGCATTAATTTCACTTAATGGAATATGTGAGCAATCGTCTTGATTAAGTA 1522

Db 705 CAGAAATAGATTCATATCTCTCTCATTTGGAATATGTTTCAGATCAATCACTAATTAATTTGAG 764

QY 1523 AAGACGGACGATGTCTGAGGAAAAGTGGGGTTAATGTTCAAGATGATCTTCGAGAGACG 1582

Db 765 AAGGCTTCCGATGTTCTTCCGAAACCTTGCAATCAGGTGCAATGATCTCTAGGGGTGA 824

QY 1583 TCAAGGTTAACATGTGCTGATAGTCCATGATATAGCGATGCAAGGCACTGTAGAACCC 1642

Db 825 TCTAAGGTGAAACATCTCATTTGTTGTAATATGACAGTAGAGCGAGATGTGTAGGGCT 884

QY 1643 CTTCATCAAGCGCTGTTGGAAGAGAGTGC 1672

Db 885 CTCACCTTAGCTTCTTGAAGATGAGCTC 914

Search completed: March 23, 2004, 13:14:24
Job time : 711 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 07:05:34 ; Search time 5264 Seconds

(without alignments)
11260.720 Million cell updates/sec

Title: US-09-890-813-5

Perfect score: 1985

Sequence: 1 gcacacagacagacagagag.....aataacattcaatct 1985

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

EST:*

1: em_estda:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vtl:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_tod:*
26: em_gss_png:*
27: em_gss_vtl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	795.2	40.1	3000	11	AY109464 Zea mays
2	635.4	32.0	682	14	CD440698 EL01N0558
3	584.2	29.4	756	14	CB666188 OS0NED12P
4	578	29.1	665	14	CD441517 EL01N0558

5	513.4	25.9	549	14	CA250542
6	477.4	24.1	512	10	AM285751
7	468.4	23.6	502	10	AM285752
8	455.4	22.9	1103	14	CK206083
9	445.2	22.4	1101	14	CK206100
10	444	22.4	827	14	CB679618
11	438.6	22.1	792	13	BQ506340
12	437.6	22.0	764	14	CF449571
13	436.2	22.0	548	10	BQ63377
14	434.2	21.9	727	13	BQ986067
15	431	21.7	607	13	BQ986067
16	425.8	21.5	822	14	CB623763
17	420.2	21.2	673	9	AI901511
18	419.8	21.1	630	14	CA931918
19	418.8	21.1	680	14	CA932103
20	417.8	21.0	653	28	BZ328792
21	417.8	21.0	839	29	CG376871
22	416.4	21.0	911	29	CG200786
23	416.4	21.0	960	28	CG413956
24	416.2	21.0	836	29	CG151954
25	412.2	20.8	692	12	BG441599
26	407.6	20.5	733	13	BQ993989
27	402.2	20.3	690	12	BQ595665
28	398.6	20.1	658	12	BU301147
29	395.6	19.9	689	13	CA064976
30	394.4	19.9	660	14	CD054614
31	388.2	19.6	913	28	CG373518
32	383.4	19.3	891	29	CG322733
33	373.4	18.8	895	29	CG322743
34	366.2	18.4	605	14	CF093872
35	364.8	18.4	697	14	CA219909
36	364.6	18.4	559	14	CF098434
37	364	18.3	706	14	CF395448
38	360.4	18.2	645	14	CF322694
39	357.6	18.0	645	10	BE249753
40	356.6	18.0	691	13	BQ148318
41	356.2	17.9	452	12	BG158155
42	356	17.9	356	9	A1629668
43	355.6	17.7	755	13	BU005561
44	352	17.7	774	14	CB619049
45	349	17.6	618	10	BE458252

ALIGNMENTS

RESULT 1	AY109464	3000 bp	MRNA	linear	HTC 17-OCT-2002
LOCUS	AY109464				
DEFINITION	Zea mays CL754_1 mRNA sequence.				
ACCESSION	AY109464				
VERSION	AY109464.1	GI:21213193			
KEYWORDS	HTC.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Bukayeva; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD				
AUTHORS	Hailey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.				
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of				
JOURNAL	Overgo Probes				
REFERENCE	Unpublished (2002)				
AUTHORS	2 (bases 1 to 3000)				
TITLE	Coe, E.H.				
JOURNAL	Direct Submission				
COMMENT	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA				
	If you are interested in getting corresponding physical clones, these are publicly available from ZMDB and may be found by BLAST searching at MSU, maizemap.org , ZMDB, www.zmdb.iastate.edu , TIGR, www.tigr.org , or NCBI, www.ncbi.nlm.nih.gov . When the source of the				

Best Local Similarity 97.7%; Pred. No. 3.7e-146;
Matches 667; Conservative 0; Mismatches 11; Indels 5; Gaps 2;

QY 345 TCATCTGACGTTCCCGAGAGAGCGCCGCTGCTCTCTCCATGGGAAAACCA 404
Db 1 TCATCTGACGTTCCCGAGAGAGCGCCGCTGCTCTCTCCATGGGAAAACCA 60
QY 405 CCAACAACTG 464
Db 61 CCAACAACTG 120
QY 465 AATCGAAGAGTGAATATGTCATAAGCTCATATCAAGCGTGTGATGAATCTGAC 524
Db 121 AATCGAAGAGTGAATATGTCATAAGCTCATATCAAGCGTGTGATGAATCTGAC 180
QY 525 TTCCAAAGATCTCTATACAG---CCTTATGAACTGGACCACTATGAAAGTATC 580
Db 181 TTCCAAAG-ATCTGTAATACAGACATGTAAGATGTAAGCACTATGAAAGTATC 239
QY 581 GCTATGTAAGAGCTGACGCTAGAGACAGTACTACTCTGTTTCAATTTGAGATGC 640
Db 240 GCTATGTAAGAGCTGACGCTAGAGACAGTACTACTCTGTTTCAATTTGAGATGC 299
QY 641 ATGTCCACAGATTTTTTCTGTTTATTTGAACAAATTCGTCAAGCGACGATAT 700
Db 300 ATGTCCACAGATTTTTTCTGTTTATTTGAACAAATTCATGTCAAGCGACGATAT 359
QY 701 GAGGCAATTTGATTTGTTTCTTATTAACAGATGTAATTTGTAATGCGATCTTAGAA 760
Db 360 GAGGCAATTTGATTTGTTTCTTATTAACAGATGTAATTTGTAATGCGATCTTAGAA 419
QY 761 GCAACCTATCTCTGCTGTGGAAGAGATTCATGCGGATGTAATACAGATCCAGATTA 820
Db 420 GCAACCTATCTCTGCTGTGGAAGAGATTCATGCGGATGTAATACAGATCCAGATTA 479
QY 821 CCGTGTGTAAGGCTGCTGCTGGAAGAGGCTGGAATCTGCTGTAATCTCTTAGAGC 880
Db 480 CCGTGTGTAAGGCTGCTGCTGGAAGAGGCTGGAATCTGCTGTAATCTCTTAGAGC 539
QY 881 CGAGTGTGATGATCTGCTGTAACCAATGTAAGCCTTGGGATGAGAGAAAT 940
Db 540 CGAGTGTGATGATCTGCTGTAACCAATGTAAGCCTTGGGATGAGAGAAAT 599
QY 941 CAGGATGGAAGATGTTGATGCTGTAATCTCTGATCCAAATCTTACCCAGATGCA 1000
Db 600 CAGGATGGAAGATGTTGATGCTGTAATCTCTGATCCAAATCTTACCCAGATGCA 659
QY 1001 AAGACTGTCCATATTAACATT 1023
Db 660 AAGACTGTCCATATTAACATT 682

RESULT 3
CB66188 756 bp mRNA linear EST 09-APR-2003
LOCUS OSJNE12P16.F OSJNE12P16 5', mRNA sequence.
DEFINITION
ACCESSION
CB66188.1 GI:29669913
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
EST.
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretidoideae; Oryzoideae; Oryza.
1 (bases 1 to 756)
Janda, D., Dean, R., Soderlund, C., Wang, R. and Wang, G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute

University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: <http://genome.arizona.edu>
PCR Primers
FORWARD: gta aac cga cga cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 12 row: P column: 16
Seq primer: gta aac cga cga cca gtc.
Location/Qualifiers
1..756
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cui_type="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNE12P16"
/tissue_type="leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNE12"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (9240-1)"

ORIGIN

Query Match 29.4%; Score 584.2; DB 14; Length 756;
Best Local Similarity 88.5%; Pred. No. 1.9e-133;
Matches 647; Conservative 0; Mismatches 78; Indels 6; Gaps 1;

QY 540 ATACAGCCTTTATGACACGAGCAATTTGAAGATGCTATGATGAAGAGCTGA 599
Db 32 ACCTATGCTGATGATGACGAGAGCTCTTGAAGATTTCAATGATGAAGAGCTGA 91
QY 600 CGCCTAGGACAGTGAATCTGTTTCAATTTGGAATGATGCAACAGATTTTTT 659
Db 92 CACTTGAGGACACTGACCTGTTTCAATTTGGAATGATGATGCAACAGATTTTTT 151
QY 660 CTGCTATTTTGAACAAATTCGTCAAGGACGCGATATGACGATTTGATTTG 719
Db 152 CTGCTATTTTGAACAAATTCGTCAAGGACGCGATATGACGATTTGATTTG 211
QY 720 TCATTAACAATGATTTGTTGTAATGCGATATCTTGAAGCAACTATCTGTTG 779
Db 212 TCATTAACAATGATTTGTTGTAATGCGATATCTTGAAGCAACTATCTGTTG 271
QY 780 CGAAGAGCTTCAATGCGGACCTGATACAGATCCAGCATCTGTTGTTGCTCC 839
Db 272 CGAAGAGCTTCAATGCGGACCTGATACAGATCCAGCATCTGTTGTTGCTCC 331
QY 840 TTGGGAAGGCTGGAATTCGTGCTGTAATCTTATGAGCGAGTGTGACTGA 899
Db 332 TTGGGAAGGCTGGAATTCGTGCTGTAATCTTATGAGCGAGTGTGACTGA 391
QY 900 CTGCTAACCAATGCTTAAGCTTGGGACCTGAGAGAAATTCAGTATGAAATGTTG 959
Db 392 CTGCTAACCAATGCTTAAGCTTGGGACCTGAGAGAAATTCAGTATGAAATGTTG 451
QY 960 ATGCTGATCTCTCTGTAATTCGAATATCTTACCCAGATGCAAGACTGTCATCTTA 1019
Db 452 ATGCTGATCTCTCTGTAATTCGAATATCTTACCCAGATGCAAGACTGTCATCTTA 511
QY 1020 CATTTGAAGAGGACAGAACTGCTTATTTTGTCTGAGTTTGCATCOAATGCA 1079
Db 512 CATTTGAAGAGGACAGAACTGCTTATTTTGTGCTGAGTTTGCATCOAATGCA 571
QY 1080 TGAGACCTGCTGAGAGAGTGAATTTCCAGTATGAGTTGAAGATTCATCAACCTTAAG 1139
Db 572 TGAGACCTGCTGAGAGAGTGAATTTCCAGTATGAGTTGAAGATTCATCAACCTTAAG 631
QY 1140 CTCAGGCAACCTTATTAACAGCAAGAGATGATGATGATGATGATGATGATGATGAT 1199
Db 632 CTCAGGCAACCTTATTAACAGCAAGAGATGATGATGATGATGATGATGATGATGAT 685

QY 1200 GCATAGTCTCAAGTCAATATGCTATATGTCAGATTGTGAGCACTGGATGCTTGGTC 1259
 DB 686 GCATAGTCTCAAGTCAATATGCTATATGTCAGATTGTGAGCACTGGATGCTTGGTC 745
 QY 1260 AGTATGCTTTT 1270
 DB 746 AATTGCTTTT 756
 RESULT 4
 CD441517 665 bp mRNA linear EST 03-JUN-2003
 LOCUS EL01N0558F11.9 Endosperm_5 Zea mays cDNA, mRNA sequence.
 ACCESSION CD441517
 VERSION CD441517.1 GI:31357160
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 665)
 REFERENCE Lai J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and
 AUTHORS Messing, J.
 TITLE Sequencing of the maize endosperm ESTs
 JOURNAL Unpublished (2002)
 COMMENT Contact: Lai, Jinsheng
 Dr. Joachim Messing's Lab
 Waksman Institute, Rutgers University
 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
 Tel: 732-445-3801
 Fax: 732-445-5735
 Email: jlai@waksman.rutgers.edu
 Seg primer: 77
 FEATURES
 source Location/Qualifiers
 1..665
 /organism="Zea mays"
 /mol_type="mRNA"
 /culivar="W22"
 /db_xref="taxon:4577"
 /issue_type="Endosperm of 7-23DAP"
 /clone_lib="Endosperm_5"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"
 ORIGIN
 Query Match 29.1%; Score 578; DB 14; Length 665;
 Best Local Similarity 94.1%; Pred. No. 6.1e-132;
 Matches 637; Conservative 0; Mismatches 26; Indels 14; Gaps 3;
 QY 1166 AGAGACATGATTAAGGTCTGTTGTTACTACTAGCATATGCTCAAGTCAATGTCAT 1225
 DB 665 AGACAAAGACATGATTAAGGTCTGTTGTTACTACTAGCATATGCTCAAGTCAATGTCAT 606
 QY 1226 AACTTGAGATGTGAGCACTGGATGCTGCTGATGCTGATGCTTCTGGCAAGGCTATCA 1285
 DB 605 AACTTGAGATGTGAGCACTGGATGCTGCTGATGCTGATGCTTCTGGCAAGGCTATTT 546
 QY 1286 GGTATTTGCTATATTAAGATCTATGATATGCTGATGCTGTTGCTACCAAGTGAAGTT 1345
 DB 545 GGTATTTGCTATATTAAGATCTATGATATGCTGATGCTGTTGCTACCAAGTGAAGTT 492
 QY 1346 AGTGTCTCTGCTGCTATGATCCATCAAGATCTGGAGTAGGAGCACTGATACAGAGCA 1405
 DB 491 AGTGTCTCTGCTGCTATGATCCATCAAGATCTGGAGTAGGAGCACTGATACAGAGCA 435
 QY 1406 AGTGAATGACCATGATGATGTAAGAGCTTGAAGAAATAGCAATTTGTTGTTCTACTTCA 1465
 DB 434 ---GAATGACCATGATGATGTAAGAGCTTGAAGAAATAGCAATTTGTTGTTCTACTTCA 378
 QY 1466 CAAGAGGCGCATATTTCACTTATCGGAATGTGGAGCAATGCTCTCTATCTGAAAG 1525

DB 377 CAGAGGCGCATATTTCACTTATCGGAATGTGGAGCAATGCTCTCTATCTGAAAG 318
 QY 1526 ACCGGACGTGTGCTGAGGAAAAGTGGGTTATGTTCAATGATCTCGAAGAGCTCA 1585
 DB 317 ACCGGACGTGTGCTGAGGAAAAGTGGGTTATGTTCAATGATCTCGAAGAGCTCA 258
 QY 1586 AAGGTTAAATGTCGTGATATGTCATGATGATGATGATGATGATGATGATGATGATGAT 1645
 DB 257 AAGGTTAAATGTCGTGATATGTCATGATGATGATGATGATGATGATGATGATGATGAT 198
 QY 1646 CATCAGGCTTCTTTGAAGACGATGCTATATCAAGTGGAGCGGAACTACTGCTG 1705
 DB 197 CATCAGGCTTCTTTGAAGACGATGCTATATCAAGTGGAGCGGAACTACTGCTG 138
 QY 1706 GGCCTGATCAAGTGGCTTTGCTGGGTCGAGGCGGTATATGTTAGATCCCACTCG 1765
 DB 137 GGCCTGATCAAGTGGCTTTGCTGGGTCGAGGCGGTATATGTTAGATCCCACTCG 78
 QY 1766 CCTCATGAAACGCGATGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1823
 DB 77 CCTCATGAAACGCGATGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 18
 QY 1824 TCTTCCAGTCTCTGAGC 1840
 DB 17 TCTTCCAGTCTCTGAGC 1
 RESULT 5
 CA250542 549 bp mRNA linear EST 25-SEP-2003
 LOCUS SCRUF1118F04.g F11 Saccharum officinarum cDNA clone SCRUF1118F04
 DEFINITION 5', mRNA sequence.
 ACCESSION CA250542
 VERSION CA250542.1 GI:35334306
 KEYWORDS EST.
 SOURCE Saccharum officinarum
 ORGANISM Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum.
 1 (bases 1 to 549)
 REFERENCE Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
 AUTHORS The libraries that made SUCSEST
 JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 COMMENT Contact: Arruda, P.
 Centro de Biologia Molecular e Engenhariaia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parnuda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bcccenter.fcav.unesp.br
 Plate: 118 row: F column: 04
 Seg primer: 17 Promoter Primer.
 FEATURES
 source Location/Qualifiers
 1..549
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCRUF1118F04"
 /lab_host="DH10B"
 /clone_lib="F11"
 /note="Organ: Inflorescence at beginning of development
 (1cm-long). Vector: pSPORT1; Site_1: SalI; Site_2: NotI;
 An unidirectional cDNA library generated from
 Inflorescence at beginning of development (1cm-long).
 cDNA was prepared from polyA+ mRNA using Superscript.
 plasmid System Kit (Invitrogen). The double-strand cDNAs
 were fractionated in a sepharose CL-2B 40cm-columns and
 fragments sizing between 0.8 and 1.5 kb were
 directionally cloned into the vector. Details of each

source of RNA and library construction can be obtained at
http://succest.lad.ic.unictamp.br/public"

ORIGIN

Query Match 25.9%; Score 513.4; DB 14; Length 549;
Best Local Similarity 96.0%; Pred. No. 5.7e-116;
Matches 526; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 576 GATGCGCTATGATGAAGAGCTGACGCCCTAGAGACAGGACTGATCTGTTTCAATTTGAG 635
DB 1 GATGCGCTATGATGAAGAGCTGACGCCCTAGAGACAGGACTGATCTGTTTCAATTTGAG 60
QY 636 AATGCATGTCCACAGAGATTTTCTGCTTATTTGAACAAATTCGTCAAGCCAGCC 695
DB 61 AATGCATGTCCACAGAGATTTTCTGCTTATTTGAACAAATTCGTCAAGCCAGCC 120
QY 696 AGATATACCATTTGATATTTGTTTCACTACAGATGAAATTTGATATGCGGATATCT 755
DB 121 AGATATACCATTTGATATTTGTTTCACTACAGATGAAATTTGATATGCGGATATCT 180
QY 756 TAGAAGCAACTATCCTGCTGTTGGAGAGAGACTTCATGGGACTGATACAGATCCAG 815
DB 181 TAGAAGCAACTATCCTGCTGTTGGAGAGAGACTTCATGGGACTGATACAGATCCAG 240
QY 816 CGATACCTGTTGTTTACTGCGTTCTTGGAGAGGCGCTGGAATCTGCTGTAATCTT 875
DB 241 CGATACCTGTTGTTTACTGCGTTCTTGGAGAGGAGAAATCAGATGCTGTTACTACTT 300
QY 876 TAGGCGAGGTGATGATGACTGACTGCTACCAACCATTTGTTAAAGCTTGGAGACTGAG 935
DB 301 TAGGCGAGGTGATGATGACTGACTGCTACCAACCATTTGTTAAAGCTTGGAGACTGAG 360
QY 936 AAATTCAGGTATGAGAAAGATGTTGATGTTGTTACTTCTGATCCAAATTCACCCAC 995
DB 361 AAATTCAGGTATGAGAAAGATGTTGATGTTGTTACTTCTGATCCAAATTCACCCAC 420
QY 996 ATCCAAAGACTGTTCCATCTTACATTTGAAGAGGCCACAGAACTTCTATTTGGTG 1055
DB 421 ATCCAAAGACTGTTCCATCTTACATTTGAAGAGGCCACAGAACTTCTATTTGGTG 480
QY 1056 CTCAGGTTTGCATCCCAATGATGAGACTGCTAGAGAGAGTATATTCAGTTAGGG 1115
DB 481 CTCAGGTTTGCATCCCAATGATGAGACTGCTAGAGAGAGTATATTCAGTTAGGG 540
QY 1116 TTAAGAT 1123
DB 541 TTAAGAT 548

RESULT 6
AM285751 512 bp mRNA linear EST 19-JUL-2000
LOCUS LG1.223.B06.b1_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
DEFINITION
ACCESSION AM285751
VERSION AM285751
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor (sorghum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Sorghum.
1 (bases 1 to 512)
Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C. and Pratt, L.H.
An EST database from Sorghum: light-grown seedlings
Unpublished (2000)
On Jan 6, 2000 this sequence version replaced gi:6675595.
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210

Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 506
POLYA=No.

FEATURES
source location/Qualifiers
1..512
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_lib="Light Grown 1 (LG1)"
/note="Organ: 10- to 14-day-old light-grown (greenhouse)
seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2:
EcoRI; The library was made from poly-A RNA in the cloning
vector lambda Zap II. Clones to be sequenced were
prepared by mass excision."

ORIGIN

Query Match 24.1%; Score 477.4; DB 10; Length 512;
Best Local Similarity 96.5%; Pred. No. 4.6e-107;
Matches 500; Conservative 0; Mismatches 12; Indels 6; Gaps 1;

QY 751 TATCTTAGAGCACTTCTGCTGTTGCCAAGAGACTTCATGGGACTGATACAGGA 810
DB 1 TATCTTAGAGCACTTCTGCTGTTGCCAAGAGACTTCATGGGACTGATACAGGA 60
QY 811 TCCAGGATACCTGTTGTTTACTGCGTTCTTGGAGAGGCGCTGGAATCTGCTGTAC 870
DB 61 TCCAGGATACCTGTTGTTTACTGCGTTCTTGGAGAGGCGCTGGAATCTGCTGTAC 120
QY 871 TACTTAGGCGAGGTGATGATGACTGCTGCTACCAACCATTTGTTAAAGCTTGGAGCT 930
DB 121 TACTTAGGCGAGGTGATGATGACTGCTGCTACCAACCATTTGTTAAAGCTTGGAGCT 180
QY 931 GAGAGAAATTCAGTATGAGAAAGATGTTGATGTTGTTACTTCTGATCCAAATATCTA 990
DB 181 GAGAGAAATTCAGTATGAGAAAGATGTTGATGTTGTTACTTCTGATCCAAATATCTA 240
QY 991 CCCACATGAAAGACTGTTCCATCTTACATTTGAAGAGGCCACAGAACTTGTATTT 1050
DB 241 CCCACATGAAAGACTGTTCCATCTTACATTTGAAGAGGCCACAGAACTTGTATTT 300
QY 1051 TGGTGCTCAGTTTGCATCCCAATGATGAGACTGCTAGAGAGTATATTCAGT 1110
DB 301 TGGTGCTCAGTTTGCATCCCAATGATGAGACTGCTAGAGAGTATATTCAGT 360
QY 1111 TAGGTTTAAAGATTCATACAACTTAAGCTCCAGGACCTTATTTCCAGACAAAGA 1170
DB 361 TAGGTTTAAAGATTCATACAACTTAAGCTCCAGGACCTTATTTCCAGACAAAGA 420
QY 1171 CATGATTAAGTCTGTTGTTTACTACTAGATAGTGTCAAGTCAATGTCACTATGTT 1230
DB 421 CATGATTA-----GTTTACTACTAGATAGTGTCAAGTCAATGTCACTATGTT 474
QY 1231 GGCATTTGAGCACTCGGATGCTTGTGTCAGTATGTT 1268
DB 475 GGCATTTGAGCACTCGGATGCTTGTGTCAGTATGTT 512

RESULT 7
AM285752 502 bp mRNA linear EST 19-JUL-2000
LOCUS LG1.223.B07.b1_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
DEFINITION
ACCESSION AM285752
VERSION AM285752
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor (sorghum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoidae; Andropogoneae; Sorghum.
 1 (bases 1 to 502)
 Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C. and Pratt, L.H.
 An EST database from Sorghum: light-grown seedlings
 Unpublished (2000)
 On Jan 6, 2000 this sequence version replaced gi:675596.
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology,
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmp@prattuga.edu
 Sequences have been trimmed to exclude polyA, vector and regions
 below phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: JEN REV
 High quality sequence stop: 353
 POLYA=No.

FEATURES
 source
 1..502
 Location/Qualifiers
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /db_xref="taxon:4558"
 /clone_lib="Light Grown 1 (LGI)"
 /note="Organ: 10- to 14-day-old light-grown (greenhouse)
 seedlings; Vector: Lambda Zap; Site 1: XhoI; Site 2:
 EcoRI; The library was made from poly-A RNA in the cloning
 vector lambda Zap II. Clones to be sequenced were
 prepared by mass excision."

Query Match 23.6%; Score 468.4; DB 10; Length 502;
 Best Local Similarity 96.7%; Pred. No. 7.8e-105;
 Matches 491; Conservative 0; Mismatches 11; Indels 6; Gaps 1;

QY 751 TATCTTNGAAGCACTATCCCTGCTTTCGGAAGACCTTGAGGAGCTGGATACGGA 810
 Db 1 TATCTTNGAAGCACTATCCCTGCTTTCGGAAGACCTTGAGGAGCTGGATACGGA 60
 QY 811 TCAGAGGATACCTGTTGTTACTGGGTTCCCTGGAAGGAGGCTGGAATCTGTGCTGTAAC 870
 Db 61 TCAGAGGATACCTGTTGTTACTGGGTTCCCTGGAAGGAGGCTGGAATCTGTGCTGTAAC 120
 QY 871 TACTTTAGCCGAGGCTGTAAGTGAATGCTGTAACCATTTGGTAAAGCTTGGGACT 930
 Db 121 TACTTTAGCCGAGGCTGTAAGTGAATGCTGTAACCATTTGGTAAAGCTTGGGACT 180
 QY 931 GAGAGAAATCAGGTATGGAAGATGTTAGTGTACTTGTGATCCCAATATCTA 990
 Db 181 GAGAGAAATCAGGTATGGAAGATGTTAGTGTACTTGTGATCCCAATATCTA 240
 QY 991 CCGACATGCAAAAGCTGTTCCATCTTAACATTGGAAGGCCACAGAACTGCTTATTT 1050
 Db 241 CCGACATGCAAAAGCTGTTCCATCTTAACATTGGAAGGCCACAGAACTGCTTATTT 300
 QY 1051 TGGTGCTCAGGTTTGTGATCCATGATGATGAGACCTGCTAGAGAAGGTGATATCCAGT 1110
 Db 301 TGGTGCTCAGGTTTGTGATCCATGATGATGAGACCTGCTAGAGAAGGTGATATCCAGT 360
 QY 1111 TAGGGTTAAGAAATTCATCAACCCCTTAAGCTCCAGGACCCCTTATTTACCAAGAAAGGA 1170
 Db 361 TAGGGTTAAGAAATTCATCAACCCCTTAAGCTCCAGGACCCCTTATTTACCAAGAAAGGA 420
 QY 1171 CATGATTAAGGCTGTTGTTGTTACTACTAGCATAGTCAAGTCAATGTACTATATTT 1230
 Db 421 CATGATTA-----GGTTGTTACTACTAGCATAGTCAAGTCAATGTACTATATTT 474
 QY 1231 GGACATTGTGAGCACTGGATGCTTGGT 1258
 Db 475 GGATATCGTGTGACCTGGATGCTTGGT 502

RESULT 8
 CK206083 1103 bp TRNA linear EST 08-DEC-2003
 LOCUS
 DEFINITION
 FGAS017659 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
 aestivum cDNA, mRNA sequence.
 ACCESSION
 CK206083
 VERSION
 CK206083.1 GI:39568473
 KEYWORDS
 SOURCE
 ORGANISM
 Triticum aestivum (bread wheat)
 Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooidae; Triticeae; Triticum.
 EST.
 1 (bases 1 to 1103)
 Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,
 Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,
 Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilson, D.,
 Penhiket, C., Roach, J.L. and Sarhan, F. and Sarhan, F.
 Functional Genomics of Abiotic Stresses in Wheat and Canola Crops
 Unpublished (2003)
 Contact: Wm L Crosby
 Bioinformatics
 University of Saskatchewan, Department of Computer Science
 1C01 Engineering Building, 57 Campus Drive, Saskatoon,
 Saskatchewan, S7N 5A9, Canada
 Tel: 306 966 1769
 Fax: 306 966 2033

Email: fgas_est@cs.usask.ca
 This sequence is the direct result of the Base calling software
 phred (default parameters). It is the raw base calls. To aid in the
 identification of the high quality insert the software Lucy
 (default parameters) has been run on this sequence. Lucy identified
 the region [18,813].
 Plate: L5B06 row: D column: 14.
 Location/Qualifiers
 1..1103
 Location/Qualifiers

FEATURES
 source

/organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
 /note="Vector: pCMV.SPORT6; Crown and developmental stages
 of spike formation in wheat cultivar Norstar. 4 mRNA
 populations were combined before constructing the library.
 The first mRNA population is from 1cm crown sections after
 30 days of cold acclimation. The second is from 1cm crown
 sections after 11 days of deacclimation (before
 deacclimation plants were fully vernalized for 49 days).
 The third is from different developmental stages of spike
 formation (5 to 50mm) that still have not emerged from the
 leaf (dissection required). The last is from different
 developmental stages of spike and seed formation after
 having emerged from the leaf (visible). First strand
 synthesis in this library was done in the presence of
 methylated dCTP thereby protecting from internal cleavage
 with NotI."

ORIGIN

Query Match 22.9%; Score 455.4; DB 14; Length 1103;
 Best Local Similarity 73.5%; Pred. No. 2e-101;
 Matches 627; Conservative 0; Mismatches 208; Indels 18; Gaps 3;

QY 846 AGGCGTGAAGATCTGGTGTGAATCTTTAGGCCGAGGTGTAGTACTGCTGTA 905
 Db 26 AGGCGTGAAGATCTGGTGTGAATCTTTAGGCCGAGGTGTAGTACTGCTGTA 85
 QY 906 CAACCATTTGGTAAAGCTTGGAGTGAAGAAATTCAGGTATGGAAGATTTGATG 965
 Db 86 CAACCATTTGGTAAAGCTTGGAGTGAAGAAATTCAGGTATGGAAGATTTGATG 145
 QY 966 TACTTACTTGTGATCCAAATATCTACCAATGCAAGACCTGTTCCATCTTAACATTTG 1025
 Db 146 TGTGACGTGTGATCCAAATATTTATGCAAAATGGGTACCAAGTACCTTACTGCTTTG 205

QY 1026 AAGAGGCGACAGAACTGGTTATTTGGTGTCTGAGGTTTGGATCCAGATGATGAGAC 1085
 DB 206 ATGAGGCGAGTGAATCTGTTATTTGGTGTGACAGGTTTGGATCCCAATCGATCGAC 265
 QY 1086 CTGTGAGAGAGGTGATATCCGATTAGGTTAAAGATTCATCAACCACTTAAGCTGAC 1145
 DB 266 CAGTGAAGGAGGATGATCCGATTCGATGAGAGTCAATATACGCTCATGACCTG 325
 QY 1146 GCACCTTATTAACGACCAAGAGACATGATTAAGTCTGTTGATTACTAATGACATG 1205
 DB 326 GCACGTGATCACTAAACAAAGAGATATGGCGAAG-----CATATTAACGACATG 379
 QY 1206 TGCTCAAGTCAATATGCTATGTTGGACATTTGTGACCTGGATGCTTGGTCAATG 1265
 DB 380 TCCTGAATCAAAATATTAACATCTGATATATGAGCAAGATGCTCGGACGATG 439
 QY 1266 GTTTCTGCAAGGGGTATCAGTATTTGCTATTTGAAGATCAATGATATCTGTGATT 1325
 DB 440 GCTTTCTAGCTTAAGTCTTCTCAATAT-----TTGAAGATTTGGTATCTGTGATT 493
 QY 1326 GTGTGCTACCAAGTGAATGATTTGTGTCTGATTCATTCATCAAAAGATCTGAGTA 1385
 DB 494 CTGTGCTACTAGTGAAGTGAAGATCAATTCATGACATGATCCATCAAAACGTGAGTGC 553
 QY 1386 GGGAACTGATACAGACAGGCAAGTGAATGACCATGATGATGAAGCTTGAGAAATAG 1445
 DB 554 GTGAATGATCCAGCAG-----GAGCTGATGATGATGATGAAGCTTGAAAGATTG 607
 QY 1446 CAATGTTGCTGCTACTTACCTGACAGAGGCGGATTAATTCATTAATGCAAAATGTGACAT 1505
 DB 608 CAGTTGTTCACTCTCTCAACAGACAGATCAATCATCTCCCTGATGAGGAACTGCGAGAT 667
 QY 1506 CGTCTCTGATACATAAGAAAGAGGGAAGTGTGCTGAGGAAAGTGGGGTTATGTTTACA 1565
 DB 668 CGTCTCTGATTTCTTGAGAGGCGCTGCAACCTGCTTACGCAAAATGTTTATGTCAGA 727
 QY 1566 TGATCTGCAAGAGCGCTCAAAAGTTAAACATGTGCTGATGATCCATGATGACATGCAA 1625
 DB 728 TGATTTGCAAGGGGCGCTTCAAGGTGAACATCTCTGTTGATGATGACAGCGAGGGA 787
 QY 1626 AGGACCTGCTGAGAAAGCCCTTCACTGAGGCTTCTTTGAAGAGATGCTCATACAAAGTCG 1685
 DB 788 AGCAGTGCAGAGAGCCCTTCACTGAGGCTTCTTGAAGAGGTTTCTGTGAGAGTGC 847
 QY 1686 AAGCGAGAACT 1698
 DB 848 AGGAGCGGACCT 860
 RESULT 9
 LOCUS CK206100
 DEFINITION Triticum aestivum FGAS: Library 5 GATE 7 Triticum
 ACCESSION CK206100
 VERSION CK206100.1 GI:39568490
 KEYWORDS EST
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 1101)
 Allard, F., Crosby, M.L., Danyluk, J., Budes, F., Frick, M., Gaudet, D.,
 Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,
 Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,
 Peniket, C., Roach, J.L., and Sarhan, F.
 Functional Genomics of Abiotic Stress in Wheat and Canola Crops
 Unpublished (2003)
 CONTACT: Wm L Crosby
 Bioinformatics
 University of Saskatchewan, Department of Computer Science
 1C101 Engineering Building, 57 Campus Drive, Saskatoon,

Saskatoon, S7N 5A9, Canada
 Tel: 306 966 1769
 Fax: 306 966 2033
 Email: fgas.est@cs.usask.ca
 This sequence is the direct result of the Base calling software
 phred (default parameters). It is the raw base calls. To aid in the
 identification of the high quality insert the software Lucy
 (default parameters) has been run on this sequence. Lucy identified
 the region (1,853).
 Plate: L5B06 row: E column: 14.
 Location/Qualifiers
 1..1101
 /organism="Triticum aestivum"
 /mol_type="RNA"
 /db_xref="taxon:4565"
 /clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
 /note="vector: pCMV.SPORT6; Crown and developmental stages
 of spike formation in wheat cultivar Norstar. 4 mRNA
 populations were combined before constructing the library.
 The first mRNA population is from 1cm crown sections after
 30 days of cold acclimation. The second is from 1cm crown
 sections after 11 days of deacclimation (before
 deacclimation plants were fully vernalized for 49 days).
 The third is from different developmental stages of spike
 formation (5 to 50mm) that still have not emerged from the
 leaf (dissection required). The last is from different
 developmental stages of spike and seed formation after
 having emerged from the leaf (visible). First strand
 synthesis in this library was done in the presence of
 methylated dCTP thereby protecting from internal cleavage
 with NotI."

ORIGIN

Query Match 22.4%; Score 445.2; DB 14; Length 1101;
 Best Local Similarity 73.4%; Pred. No. 6,6e-99;
 Matches 629; Conservative 0; Mismatches 209; Indels 19; Gaps 4;
 QY 846 AAGGCTGGAATCTGTGCTGTAATCTTAAAGCCGAGGTGATGACTTGAAGCTGA 905
 DB 23 AAGGCTGGAATCAATCCGCGGTACACAGTTACGAAGGGGGGTAGTACTTGAAGCTGA 82
 QY 906 CAACATTTGGCAAAAGCTTGGGTTAAGAGAAATCAGATTGGAAGATGATGATG 142
 DB 83 CAACATTTGGCAAAAGCTTGGGTTAAGAGAAATCAGATTGGAAGATGATGATG 142
 QY 966 TACTTAATTGTATCCAAATATCTACCAATGCAAGACTGTTTCCATCTTAACATTG 1025
 DB 143 TGTGACGTGTATCAATATTTATGCAAAAGCGGTACAGTACCTTACTGACTTTG 202
 QY 1026 AAGAGCCACAGAACTTGTATTTTGTGCTCAGGTTTGCATCCAAATCGATGAGAC 1085
 DB 203 ATGAGGCAAGTGAATCTGTTATTTGGTGAAGATTTGATCCCAATCGATGAGAC 262
 QY 1086 CTGCTAGAGAGGTGATATTCAGTATGAGGTTAAGATTCATCAACCTTAAGCTCGAG 1145
 DB 263 CAGCTAGGGAAGGTGATCCAGTTCAGGTGAAGATCAATATTAACGTCATGACCTG 322
 QY 1146 GCACCTTATTAACGACCAAGAGATGAGATTAAGTCTGTTTACTAATGACATG 1205
 DB 323 GCACGTGATCACTAAACAAAGATGATGCGAAG-----CATATTAACGACATG 376
 QY 1206 TGCTCAAGTCAATATGCTATGTTGGACATTTGAGAGCTGAGTCTGTTGATGATG 1265
 DB 377 TCCTGAATTAATATTAACATGCTGATATGATGAGCAAGATGCTCGGACGATG 436
 QY 1266 GTTTTCTGCAAGGGATAGAGTATTTGTTATTTGAAGTATGATATCTGTGATT 1325
 DB 437 GCTTTCTAGTGAAGGCTTCTCAATAT-----TTGAAGATTTGGGATCTCTGTGATT 490
 QY 1326 GTGTGCTACCAAGTGAATGATGTTTCTGTGCTCACTGATCCATCAAAAGATCTGAGTA 1385
 DB 491 CTGTGCTACTAGTGAAGTGAAGATCAATATGATGACATGATCCATCAAAACGTGAGTGC 550

QY 1386 GGGAGACTGATACAGCAGGAGAGTAACTTGACCATGTAGTGAAGAGCTTGAGAAAATAG 1445
 Db 551 GTGAATTGATCCAGCAGG-----GAGCTTGATCATGTAGTGAAGAGCTTGAGAAAATAG 604
 QY 1446 CAATTGTGTCTACTTCAAGCAGAGGCGCATTAATTCACTTATCGGAATGTGAGCAAT 1505
 Db 605 CAGTTGTTCACTTCTTCAAGCAGCAGATCATATTTCTCTATAGGAGAGCTGAGAGAT 664
 QY 1506 CGTCTGTGATCTAGAAAAGACGAGCAGTGTCTGAGAGAAAAGTGGGTTATGTTCAGA 1565
 Db 665 CGTCTGTGATCTTGAAGAGCGTTCACGTTCTACGCAAGAAATGTGTATGTTCAGA 724
 QY 1566 TGAATTCGAGAGAGGCGTCAAGGTAAATGTCGCTGATAGTCATATAGCATGCA 1625
 Db 725 TGAATTCGAGAGAGGCGTCAAGGTAAATGTCGCTGATAGTCATATAGCATGCA 784
 QY 1626 AGGAGCTCTGAGAGAGCGCTTCAATCAGCGCTTCTTGAAGACGATGTCTTATCAGAGTCG 1685
 Db 785 AGGAGCTCTGAGAGAGCGCTTCAATCAGCGCTTCTTGAAGACGATGTCTTATCAGAGTCG 843
 QY 1686 AAGCGAGAACCTACTC 1702
 Db 844 AAGAGCGAGAGCTTGAC 860

RESULT 10

LOCUS CB679618 827 bp mRNA linear EST 09-APR-2003
 DEFINITION OSJNEF03F11.f OSJNEF Oryza sativa (japonica cultivar-group) cDNA
 clone OSJNEF03F11 5', mRNA sequence.

ACCESSION CB679618 GI:29683343
 VERSION CB679618
 KEYWORDS EST.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 827)
 Jantsuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
 Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.
 Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea
 Unpublished (2003)

JOURNAL COMMENT
 Contact: Rod Wang
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu

FEATURES
 source
 1..827
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNEF03F11"
 /rissue_type="leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEF"
 /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
 XhoI; Uninfected Control"

ORIGIN

Query Match 22.4%; Score 444; Db 14; Length 827;

Best Local Similarity 73.3%; Pred. No. 1,1e-98;
 Matches 610; Conservative 0; Mismatches 212; Indels 10; Gaps 3;
 QY 406 CACACACCTTCTCT-TGCTGAGAGAGAGCAGTAGAGGTGAGATTATCATGTTCTG 464
 Db 1 CAAATACCTCTCTCTTTGGAGAGAGGTGTCACTCGGGGCCGAGAGCGTCTG 60
 QY 465 AATGAGAGGTGAGATGTGTCAAAAGCTTACATACAGAGCGGTGATGAATCTTGAC 524
 Db 61 AAATTCGAGGTGAGATGTGTCAAAAGCTTACATACAGAGCGGTGATGAATCTTGAC 120
 QY 525 TTCCAGAGATGTGTATCAAGCTTT---ATGAACTGAGCAACTATGAAGTATCG 581
 Db 121 TGGATAGATCATGTTTTCAGGTTTATTTGAAGATTTGAACAACCTTTAAGGTTG 180
 QY 582 CTATATGAAAGCTGACGCTTGAACAGTACTTCTTTTCTATTTGGAAATGCA 641
 Db 181 CTATATGAAAGCTGACGCTTGAACAGTACTTCTTTTCTATTTGGAAATGCA 240
 QY 642 TGTCACAGAGATTTTCTGCTATTTGAACAAATTCGTGCAAGCAGGAGATG 701
 Db 241 TGTCTCAAGATATTTTCTGCTATTTGAACAAATTCGTGCAAGCAGGAGATG 300
 QY 702 AGCATTGATATTTGTTTCAATTAACAATGATGATTTGATTTGCGATATCTTAAG 761
 Db 301 ATGATTGATATTTGTTTCAATTAACAATGATGATTTGATTTGCGATATCTTAAG 360
 QY 762 CAACCTATCTGCTGTTGCGAGAGACTTATGAGGAGCTGATACAGATCCAGCATAC 821
 Db 361 CCACTATCTGCTGTTGCGAGAGACTTATGAGGAGCTGATACAGATCCAGCATAC 420
 QY 822 CTGTTGTTACTGAGTCTTCTTGGAGAGGCTGGAATGTGCTGTTACTTATTTAGGC 881
 Db 421 CTATGTTACTGAGTCTTCTTGGAGAGGCTGGAATGTGCTGTTACTTATTTAGGC 480
 QY 882 GAGGTGATGATCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCT 941
 Db 481 GAGGTGATGATCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCT 540
 QY 942 AGGTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1001
 Db 541 AGGTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 1002 AGCTGTTCACTTATTAATTTGAAGGCGCACAGACTTCTTATTTGTGCTCAGG 1061
 Db 601 TACCTGTTCACTTATTAATTTGAAGGCGCACAGACTTCTTATTTGTGCTCAGG 660
 QY 1062 TTTTGCATCCCAATCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1121
 Db 661 TTTTGCATCCCAATCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 QY 1122 ATTATTAACACCTTAAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1181
 Db 721 ATTATTAACACCTTAAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 QY 1182 GTCTGTTGTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1243
 Db 781 G-----TATATTAACAGCATTTATTAAGTCAACATTTACTATGCTGGA 826

RESULT 11
 LOCUS BQ506340 792 bp mRNA linear EST 07-MAR-2003
 DEFINITION EST131755 Generation of a set of potato cDNA clones for microarray
 analyses mixed potato tissues Solanum tuberosum cDNA clone STMG371
 5' end, mRNA sequence.
 ACCESSION BQ506340
 VERSION BQ506340.2 GI:21922214
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE
AUTHORS

asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 792)
Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,
Restrepo, S., Griffiths, H., van der Have, R., Tsai, J. and
Karymcheva, S.A.
Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)
On Jun 10, 2002 this sequence version replaced gi:21365209.
Contact: Robin Buell
The Institute for Genomic Research
3712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igf.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.

FEATURES
source

Location/Qualifiers
1..792
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec or Binje"
/db_xref="taxon:4113"
/clone="STMGJ71"
/tissue_type="mixed tissues"
/lab_host="SOLAR"
/clone_lib="Generation of a set of potato cDNA clones for
microarray analyses mixed potato tissues"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating
eyes, tubers, or roots."

ORIGIN

Query Match 22.1%; Score 438.6; DB 13; Length 792;
Best Local Similarity 76.1%; Pred. No. 2.4e-97;
Matches 556; Conservative 0; Mismatches 169; Indels 6; Gaps 1;
QY 554 GAACCTGAGCACTATTTGAAGATGCTATGATGAAGAGCTAGCCAGT 613
DB 63 GAACCTGAGCACTATTTGAAGATGCTATGATGAAGAGCTAGCCAGT 122
QY 614 GACTACCTGCTGCTATTTGAGATGATGCTACACAGATTTTCTGCTATTTGAA 673
DB 123 GACTACCTGCTGCTATTTGAGATGATGCTACACAGATTTTCTGCTATTTGAA 182
QY 674 AAAATTCGTGCAAGGACGCGCATGTCAGCATTTGATTTGTTTCAATTCAGTAT 733
DB 183 AAAATTCGTGCAAGGACGCGCATGTCAGCATTTGATTTGTTTCAATTCAGTAT 242
QY 734 GAATTCGATGATCGGATATCTTGAAGCACTATTCCTGCTGTTGCAAGAGACTTAT 793
DB 243 GAATTCGATGATCGGATATCTTGAAGCACTATTCCTGCTGTTGCAAGAGACTTAT 302
QY 794 GGGGACTGATACAGATCCAGGATACCTGTTTACTGGGTTCTTGGGAAGGCTGG 853
DB 303 GGGGACTGATACAGATCCAGGATACCTGTTTACTGGGTTCTTGGGAAGGCTGG 362
QY 854 AAATCTGGGCTGTAATCTTGAAGCAAGGCTGATGATGATGATGATGATGATGAT 913
DB 363 AGAAGCTTGTCTGATACATGATGATGATGATGATGATGATGATGATGATGAT 422
QY 914 GGTAAAGCTTGGGACTGAGAAATTCAGGATGAGAAAGATGTTGATGTTACTTACT 973
DB 423 GGTAAAGCTTGGGACTGAGAAATTCAGGATGAGAAAGATGTTGATGTTACTTACT 482
QY 974 TGGATGCAATATCTTACCAATGCAAGAGCTGTTCCATCTTAACTTGAAGAGCC 1033
DB 483 TGGATGCAATATCTTACCAATGCAAGAGCTGTTCCATCTTAACTTGAAGAGCC 542
QY 1034 ACAAGACTGCTTATTTGCTGCTGATTTGATGATGATGATGATGATGATGAT 1093
DB 543 GCTAACTGATCTTGGGAGGAGGCTCTTCACTCACTCACTCACTCACTCACTCA 602

QY 1094 GAAGGTGATATTCAGATTAGGATTAAAGATTCATACACCTTAAGCTCGACCCCTT 1153
DB 603 GAGGCTGATATTCCTGTTAGGTTAAAGATTCATATCCTTAAGCTCGACCTTC 662
QY 1154 ATTACCAACAAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1213
DB 663 ATCGTTTAAACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 716
QY 1214 TCAATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1273
DB 717 AGCATGATACCATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 776
QY 1274 GCAAGGATATC 1284
DB 777 CAAGGATATC 787

RESULT 12

CF449571 764 bp mRNA linear EST 04-SEP-2003
LOCUS EST85916 normalized cDNA library of onion Allium cepa cDNA clone
DEFINITION ACACU05, mRNA sequence.
ACCESSION CF449571
VERSION CF449571.1 GI:34472273
KEYWORDS EST
SOURCE Allium cepa (onion)
ORGANISM Allium cepa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
Allium.

REFERENCE 1 (bases 1 to 764)
AUTHORS Havey, M.J., Cheung, F., Van Aken, S., Uteback, T. and Town, C.D.
TITLE Expressed Sequence Tags from a normalized library of mixed onion
tissues (Allium cepa)
JOURNAL Unpublished (2003)
COMMENT Contact: Havey MJ
Department of Horticulture
USDA-ARS and University of Wisconsin
1575 Linden Drive, Madison, WI 53706, USA
Tel: 608-262-1830
Fax: 608-262-4743
Email: mjhvey@facstaff.wisc.edu
TIGR sequence name ACACU05TR. For more information:
http://haveylab.hort.wisc.edu
Seq primer: CAG GAA ACA GCT ATG ACC.

FEATURES
source

Location/Qualifiers
1..764
/organism="Allium cepa"
/mol_type="mRNA"
/cultivar="Red Creole (bulbs), unknown (callus), Ebano &
Texas Legend (roots)"
/db_xref="taxon:4679"
/clone="ACACU05"
/tissue_type="Callus, roots, and young bulbs"
/clone_lib="normalized cDNA library of onion"
/note="Vector: pCMVSPORT6.1-cdb (Invitrogen); Site_1:
EcoRV (5'); Site_2: NotI (3'); Equal molar amounts of mRNA
from callus, roots, and young bulbs were combined to
synthesize the library. Normalization to enrich for
low-copy transcripts was performed by proprietary
techniques of Invitrogen."

ORIGIN

Query Match 22.0%; Score 437.6; DB 14; Length 764;
Best Local Similarity 74.8%; Pred. No. 4.2e-97;
Matches 579; Conservative 0; Mismatches 184; Indels 11; Gaps 2;
QY 581 GCTATGATGAAGAGCTGACGCTGACCAAGTCACTCTGTTTCAATTGAGAAATGC 640
DB 1 GCTATGATGAAGAGCTTACCTCCGTGCAAGATCTTTAGTTTCTTAAGGAGAGCC 60
QY 641 ATGTCCACCAAGATTTTCTGCTTATTTGAACAAATTCGTGCAAGCAAGCACTAT 700

[illegible]

FEATURES					
	source	location/qualifiers			
1..548					
/organism="Sorghum propinquum"					
/mol_type="mRNA"					
/db_xref="taxon:132711"					
/clone_1fb="Rhizome2 (RH122)"					
/note="Organ: Rhizomes; Vector: pBluescript II from Lambda Zap II; Site 1: XhoI; Site 2: EcoRI. The library was made from poly-A RNA in the cloning vector lambda Zap II.					
Clones to be sequenced were prepared by mass excision."					
ORIGIN					
Query Match	22.0%;	Score 436.2;	DB 10;	Length 548;	
Best Local Similarity	90.6%;	Pred. No. 7.9e-97;			
Matches 465;	Conservative 0;	Mismatched 48;	Indels 0;	Gaps 0;	
QY	4	CCAGCAGAGGACGAGCACTCAGAATGGCAATCCAGTGCATGGCGCGGCCCGCCG	63		
Dd	36	CCCCACGAGTAGAGGGCCAGAAATGGCGACCACCAATGCCATGGCGCGCACACG	95		
QY	64	CCGCCTCGTCCTCGTGATACCTCCGCGAGCCTGTGACAATGTTGAGAACCTGGCTTT	123		
Dd	96	CCGCCTCGTCCTCGTGATACCTCCGCGAGCCTCGGACATGTTCAAACAATTGCGCTGTT	155		
QY	124	CGGTACCCGAACCGGGCCCTCGCGGTGCAABAGSGTTGTCAATGGTGGTGGCCGACTCCAC	183		
Dd	156	CGATACCCGAACCGAGCTTGTCATGCCCAGSGTTGTCAATGGTGGTGGCCGACTCCAC	215		
QY	184	CAGCCGTCGGGCAAGCAAGCGAGCGGGGAGCGGCGTCTTGGGGCGCGCTGTTCTCG	243		
Dd	216	CCGCGCGTCGGGCAAGTAGGGAGCGAGCGGGAGCGGCGTCTTGGGGCGCGCTGTTCTCG	275		
QY	244	AGGGCTCGGGATGAGAGGGATTGGGGGATTCAGCTCAACGTGGTGAATGAAGTTCCGGGGGCTC	303		
Dd	276	AGGGCTCGGGATGAGGGGATTTGGGGGATTCAGCTCAACGTGGTGAATGAAGTTCCGGGGGCTC	335		
QY	304	CTCGGTCGTGTGCGCCGCGAGAGATGCTGAGGTGGCGCGGCTTCATCTTGACCTTCCCGA	363		
Dd	336	CTCAGTGGGTCGTGCGCCGCGAGAGATGCGGAGGTGGCTTCATCTTGACCTTCCCGA	395		
QY	364	GGAGGCGCCCGTCGTCCTCTGTGCGCATGGGGAATAACAACAACACTTCTCTCTGCG	423		
Dd	396	GGAGGCGCCCGTCGTCCTCTGTGCGCATGGGGAATAACAACAACACTTCTCTCTGCG	455		
QY	424	TGAGAGGAAGCAGTAGGGGTGTGAGATTATCCATGTTTCTGAATCGAAAGTGGATAT	483		
Dd	456	TGAGAGGAATGAGTAGGGGTGTGAGATTATCCATGTTTCTGAATCGAAAGTGGATAT	515		
QY	484	GGTCAAAAAGCCATATCAAGACGGTGAATGA	516		
Dd	516	GATCAAAAGCTCATATCAAGACAGGAGATGA	548		
RESULT 14					
BQ969663	727 bp	mRNA	linear	EST 21-AUG-2002	
LOCUS	clone OHB39D05,	mrna sequence.			
DEFINITION	OHB39D05.YE.ab1 OH ABCDI sunflower RH4801 Helianthus annuus cDNA				
ACCESSION	BQ969663.1	GI:22387184			
VERSION	EST:				
KEYWORDS	Helianthus annuus (common sunflower)				
SOURCE	Helianthus annuus				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; Campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus. 1 (bases 1 to 727) Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,				

Elison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
Lai, Z., Church, S., Jackson, L., and Bradford, K.
Lecture and Sunflower ESTs from the Composite Genome Project
http://comgenomics.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmor]
Department of Vegetable Crops, R.W.Michelmor Lab
University of California at Davis (UCD)
Aamundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
Singletons, see http://cgpdb.ucdavis.edu/ for details.
Plate: QHB39 row: D column: 05.

FEATURES

Source

1..727
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"
/clone="QHB39D05"
/lab_host="E.coli"
/clone_lib="QH_ABCDI sunflower RHA801"
/note="Vector: pBRCDNA51AB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=shoots environmental stress
TAG_LIB=QH ABCDI sunflower RHA801
TAG_SEQ=TCGCAACGGG"

ORIGIN

Query Match 21.9%; Score 434.2; DB 13; Length 727;

Best Local Similarity 75.6%; Pred. No. 2.9e-96;

Matches 554; Conservative 0; Mismatches 173; Indels 6; Gaps 1;

QY 578 ATGGCTATGATGAAGAGCTGAGCGCTAGACAGTACCTGTTTCATTGGAGAA 637
Db 1 ATTCTGTACTTAAGAAATTCACACACCGCAGAGATTAACCTAGCTCTGTTGGAG 60
QY 638 TGCATGTCACACAGATTTTCTGCTTAATTGAACAAATTGCTGTCAGAGCAGCGAG 697
Db 61 TGCATGTCACACAGATTTTCTGCTTAATTGAATTAATCGGTGTCAGAGCAGCGAA 120
QY 698 TATGACGCACTTATATTTGGTTTCATTACACAGATGAATTTGGTAATCGGATCTTA 757
Db 121 TATGATGCATTCGAAATTTGTTTCATTAACAAGATGATCTTACTTAATCGGATTTTGG 180
QY 758 GAAGCAACCTATCTGCTGTTGGAAAGACTTCATGGAGATGATACAGATCCAGCG 817
Db 181 GAAGCAACCTATCCAGCTGTAGCAAGACACTAAATGATGATTAAGATCTCTCA 240
QY 818 ATACCTGTTGTACTGGTTCCTTGGAAAGCGGTGAAATTCGTGCTGTACTACTTA 877
Db 241 ATTCTATTTGTACTGCTCTCTTGAAGAGGTTGAGAACTGTGCTGTACTACTCTG 300
QY 878 GAGCGAGGTGTAGTACTGCTGTACCAACCATTTGTAAAGCCTTGGAGCTGAGAGA 937
Db 301 GAGCGAGGTGTAGTACTGCTGTACCAACCATTTGTAAAGCCTTGGAGCTGAGAGA 360
QY 938 ATTGAGGTATGAAAGATTTGATGTGTACTTACTTGTATCCAAATTTTACCAAGAT 997
Db 361 ATTGAGGTATGAAAGATTTGATGTGTACTTACTTGTATCCAAATTTTATTTGCGGC 420
QY 998 GAAAGACTGTTCCATCTTAATTAATTTGAAGAGCCACAGAACTTGCTTATTTTGGTGT 1057
Db 421 GCTGACCTGTCCATTTTGAAGCTTGTGATGAGAGCAGTGAAGCTTGTATTTTGTGTC 480

QY 1058 CAGTTTTCATTCACATTCAGTATGAGACCTGCTAGAGAGGTATATTCAGTTAGGTT 1117
Db 481 CAGTCTTCATTCACATTCAGTATGAGAGCTGCTAGAGAGGTATATTCAGTTAGGTT 540
QY 1118 AAGATTCATACACCTTAAGCTTCAGGACGCTTATTTACCAAGCAAAAGACATGAT 1177
Db 541 AAAAAATGTTATACCCCTTAAGCTTCAGTACCTTAATTAACCAATCAAGAGATGAGT 600
QY 1178 AANGCTGTGTGTACTACTAGCATATGCTCAAGTCAATGCTACTATGTTGACATT 1237
Db 601 AA-----GGCAGTACTACACAGCATTTGTTGAAGTATATGATCAATTTGACATT 654
QY 1238 GTGACACTCGAGTCTGTTGATGATGTTTCTGCAAGGATATCAGATTTGCTAT 1297
Db 655 GTTAGCACTCGCATCTGCTGTCATATTGATTTCTTGCAAGGTTTTCATATTGAG 714
QY 1298 ATTGAAGATCTAT 1310
Db 715 GATTGGGTATAT 727

RESULT 15
BU988067 607 bp mRNA linear EST 22-OCT-2002
LOCUS HPI6004 HF Hordeum vulgare subsp. vulgare cDNA clone HPI6004
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BU988067
VERSION BU988067.1 GI:242339013
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Striptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 607)
Radtchuk, V., Zhang, H., Weschke, W., Potokina, E. and Webus, U.
Barley ESTs from developing seeds
Unpublished (2002)
JOURNAL Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 607 Std Error: 0.00
Plate: 16 Row: J Column: 4
Seq primer: MJREV.

COMMENT

FEATURES

Source

1..607
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="GABI:248011"
/db_xref="taxon:112509"
/clone="HF16J04"
/tissue_type="caryopsis"
/dev_stage="developing caryopsis, 16-25 DAF (days after
flowering)"
/lab_host="XL10-Gold"
/clone_lib="HF"
/note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
cDNA); Site 2: XhoI (3'-end of cDNA); developing
caryopsis, 16-25 DAF(days after flowering) Due to a cloning
artefact caused by the kit, in most cases the EcoRI site
is NOT present, as well as the EcoRI adapter used for
cloning. To excise the insert, restriction sites upstream
EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
due to the cloning system used blue/white selection for
recombinants is not 100% reliable. Average insert size is
940 bp"

ORIGIN

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 11:04:35 ; Search time 77 seconds
(without alignments)
2054.891 Million cell updates/sec

Title: US-09-890-813-6

Perfect score: 2813

Sequence: 1 MAIPVRSAAAPRLVPSITP.....QAFEDDVLSQVEARNLVG 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum March 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2809	99.9	560	AAE04353	AAE04353 Zea mays
2	2678.5	95.2	555	AAE04354	AAE04354 Zea mays
3	1835	65.2	564	AAE04359	AAE04359 Glycine m
4	1789	63.6	569	ABE93449	ABE93449 Herbicida
5	1747.5	62.1	544	AAE04236	AAE04236 Arabidops
6	1747.5	62.1	544	ABE93460	ABE93460 Herbicida
7	1728	61.4	563	AAE04215	AAE04215 Arabidops
8	1727.5	61.4	459	AAE04237	AAE04237 Arabidops
9	1708	60.7	478	AAE04216	AAE04216 Arabidops
10	1675.5	59.6	446	AAE04238	AAE04238 Arabidops
11	1675.5	59.6	555	AAE05075	AAE05075 Arabidops
12	1675.5	59.6	559	AAE05074	AAE05074 Arabidops
13	1675.5	59.6	559	ABE92174	ABE92174 Herbicida
14	1666.5	59.2	473	AAE05076	AAE05076 Arabidops
15	1656	58.9	465	AAE046217	AAE046217 Arabidops
16	1629.5	57.9	439	AAE04358	AAE04358 Wheat wri
17	16105	39.3	307	AAE18034	AAE18034 Arabidops
18	1074.5	38.2	354	AAE28330	AAE28330 Arabidops
19	1074.5	38.2	418	AAE28329	AAE28329 Arabidops
20	1054.5	37.5	269	AAE04356	AAE04356 Arabidops
21	890.5	31.7	281	AAE04355	AAE04355 Oryza sat
22	676	24.0	215	AAE18035	AAE18035 Arabidops
23	666	23.7	152	AAE04352	AAE04352 Zea mays
24	592	21.0	465	AAE06986	AAE06986 Phototrab
25	572	20.3	449	AAE1986	AAE1986 Lysine in

26	572	20.3	449	2	AAE78266	AAE78266 E. coli a
27	572	20.3	449	2	AAE85308	AAE85308 Lysine in
28	572	20.3	449	2	AAE87739	AAE87739 E. coli w
29	572	20.3	449	6	ADA16010	ADA16010 E. coli a
30	572	20.3	449	6	ABO44355	ABO44355 E. coli L
31	572	20.3	449	6	ABO44358	ABO44358 E. coli m
32	572	20.3	449	6	ADB23057	ADB23057 E. coli f
33	571	20.3	449	2	AAE69576	AAE69576 E. coli a
34	571	20.3	449	2	AAE18820	AAE18820 Aspartoki
35	569	20.2	449	2	AAE69102	AAE69102 E. coli a
36	569	20.2	449	2	AAE76953	AAE76953 Wild-type
37	569	20.2	449	2	AAE01784	AAE01784 Wild type
38	569	20.2	449	3	AAE29515	AAE29515 Escherich
39	569	20.2	449	6	ADA16015	ADA16015 E. coli a
40	568	20.2	449	6	ADA16014	ADA16014 E. coli a
41	567	20.2	449	2	AAE69577	AAE69577 E. coli a
42	566	20.1	449	2	AAE69573	AAE69573 E. coli a
43	566	20.1	449	2	AAE69574	AAE69574 E. coli a
44	566	20.1	449	2	AAE18818	AAE18818 Aspartoki
45	566	20.1	449	2	AAE18817	AAE18817 Aspartoki

ALIGNMENTS

RESULT 1
AAE04353 standard; protein; 560 AA.

ID AAE04353;

AC AAE04353;

DT 04-SEP-2001 (first entry)

XX Zea mays choic.pK002.k6 (FIS) clone aspartate kinase.

DE Zea mays choic.pK002.k6 (FIS) clone aspartate kinase.

XX Zea mays.

OS Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

XX Zea mays.

MO200146393-A2.
28-JUN-2001.
19-DEC-2000; 2000OWO-US034396.
21-DEC-1999; 99US-0172944P.
(DUPO) DU POINT DE NEMOURS & CO E. I.
FALCO SC, Famodu OO, Thorpe CJ;
WPI; 2001-418057/44.
N-PSDB; AAD08621.
Novel nucleic acid molecules encoding aspartate kinase useful in cell transformation and transgenic plant production procedures.
Claim 21; Fig 1; 64pp; English.
The present invention relates to nucleic acid molecules encoding aspartate kinase useful in cell transformation and transgenic plant production procedures. Nucleic acid encoding aspartate kinase is useful for producing a transgenic plant, by transforming a plant cell with a chimeric gene comprising nucleic acid encoding aspartate kinase operably

CC linked to a regulatory sequence and regenerating a plant from the
 CC transformed plant cell. The present sequence is cdp1c.pk010.k6 (FIS)
 CC clone Zea mays aspartate kinase
 XX

Sequence 560 AA;
 Query Match 99.9%; Score 2809; DB 4; Length 560;
 Best Local Similarity 100.0%; Pred. No. 1,36-256;
 Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIPVSAAPRRLLVPSIPASSGHVRLACFTGTGPRARGLSMVVADSTRRAKQAD 60
 DB 1 MAIPVSAAPRRLLVPSIPASSGHVRLACFTGTGPRARGLSMVVADSTRRAKQAD 60
 QY 61 GGDGVGAPVGLGMEGLGQDLSVYMKFGSSSVSAAARAEVAGLITFPEERPVYVLS 120
 DB 61 GGDGVGAPVGLGMEGLGQDLSVYMKFGSSSVSAAARAEVAGLITFPEERPVYVLS 120
 QY 121 AMGKTNNLLAGEKAVCGVIHVSEIEEMNVKSLIKTVDELGLPRSVIQDMLDELQ 180
 DB 121 AMGKTNNLLAGEKAVCGVIHVSEIEEMNVKSLIKTVDELGLPRSVIQDMLDELQ 180
 QY 181 LKGIAMMKELTPRTSDYLVSPGCMSTRIFSAYLNKIRKAROYDAPDIFGITTFEFGNA 240
 DB 181 LKGIAMMKELTPRTSDYLVSPGCMSTRIFSAYLNKIRKAROYDAPDIFGITTFEFGNA 240
 QY 241 DILEATYPAVAKRLHGDWIDOPAIPIVGTGFLGKGMKSGAVTTLGRGSDLTATTIGKALG 300
 DB 241 DILEATYPAVAKRLHGDWIDOPAIPIVGTGFLGKGMKSGAVTTLGRGSDLTATTIGKALG 300
 QY 301 LREIQWKVDVGLTCDPNIYPAKTVPYLTPEEATELAFGAQVLAHQSMRPARBGDI 360
 DB 301 LREIQWKVDVGLTCDPNIYPAKTVPYLTPEEATELAFGAQVLAHQSMRPARBGDI 360
 QY 361 VRVKSYNPKAPGTLITRQDMDXGLVLTSLVLSNVTMLDIVSTRMLGOYGFARVSG 420
 DB 361 VRVKSYNPKAPGTLITRQDMDXGLVLTSLVLSNVTMLDIVSTRMLGOYGFARVSG 420
 QY 421 ICYIEDLCISVDCVATSEVSVSLSDSKISRELIQASSELDHVVELEKIAIVRLIQ 480
 DB 421 ICYIEDLCISVDCVATSEVSVSLSDSKISRELIQASSELDHVVELEKIAIVRLIQ 480
 QY 481 RAIIISLIGNVEQSSLIIEKTGRVLRKSGVNVQMIISQASKVMNLSLIVHSDAKALVEALH 540
 DB 481 RAIIISLIGNVEQSSLIIEKTGRVLRKSGVNVQMIISQASKVMNLSLIVHSDAKALVEALH 540
 QY 541 QAFFEDDVLSCVEAENLLVG 560
 DB 541 QAFFEDDVLSCVEAENLLVG 560

RESULT 2

AAE04354
 ID AAE04354 standard; protein; 555 AA.

XX AAE04354;
 AC 04-SEP-2001 (first entry)
 DT Zea mays cdp1c.pk010.k1 (FIS) clone aspartate kinase.
 XX Zea mays.
 OS Zea mays.
 XX MO200146393-A2.
 XX 28-JUN-2001.
 XX 19-DEC-2000; 2000MO-US034396.
 XX 21-DEC-1999; 99US-0172944P.

XX (DUPO) DU POINT DE NEMOURS & CO E.I.
 PA Falco SC, Farnou OO, Thorep CJ;
 PI WPI; 2001-418057/44.
 DR N-PSDB; AAD08622.
 XX Novel nucleic acid molecules encoding aspartate kinase useful in cell
 PT transformation and transgenic plant production procedures.
 CC Claim 21; Fig 1; 64pp; English.

CC The present invention relates to nucleic acid molecules encoding
 CC aspartate kinase useful in cell transformation and transgenic plant
 CC production procedures. Nucleic acid encoding aspartate kinase is useful
 CC for producing a transgenic plant, by transforming a plant cell with a
 CC chimeric gene comprising nucleic acid encoding aspartate kinase operably
 CC linked to a regulatory sequence and regenerating a plant from the
 CC transformed plant cell. The present sequence is cdp1c.pk010.k1 (FIS)
 CC clone Zea mays aspartate kinase
 XX

Sequence 555 AA;
 Query Match 95.2%; Score 2678.5; DB 4; Length 555;
 Best Local Similarity 96.8%; Pred. No. 36-244;
 Matches 543; Conservative 1; Mismatches 10; Indels 7; Gaps 4;

QY 1 MAIPVSAAPRRLLVPSIPASSGHVRLACFTGTGPRARGLSMVVADSTRRAKQAD 60
 DB 1 MAIPVSAAPRRLLVPSIPASSGHVRLACFTGTGPRARGLSMVVADSTRRAKQAD 60
 QY 61 GGDGVGAPVGLGMEGLGQDLSVYMKFGSSSVSAAARAEVAGLITFPEERPVYVLS 120
 DB 61 GGDGVGAPVGLGMEGLGQDLSVYMKFGSSSVSAAARAEVAGLITFPEERPVYVLS 120
 QY 121 AMGKTNNLLAGEKAVCGVIHVSEIEEMNVKSLIKTVDELGLPRSVIQDMLDELQ 180
 DB 121 AMGKTNNLLAGEKAVCGVIHVSEIEEMNVKSLIKTVDELGLPRSVIQDMLDELQ 180
 QY 180 LKGIAMMKELTPRTSDYLVSPGCMSTRIFSAYLNKIRKAROYDAPDIFGITTFEFGN 239
 DB 180 LKGIAMMKELTPRTSDYLVSPGCMSTRIFSAYLNKIRKAROYDAPDIFGITTFEFGN 240
 QY 240 ADILEATYPAVAKRLHGDWIDOPAIPIVGTGFLGKGMKSGAVTTLGRGSDLTATTIGKAL 300
 DB 240 ADILEATYPAVAKRLHGDWIDOPAIPIVGTGFLGKGMKSGAVTTLGRGSDLTATTIGKAL 300
 QY 300 GIREIQWKVDVGLTCDPNIYPAKTVPYLTPEEATELAFGAQVLAHQSMRPARBGDI 359
 DB 300 GIREIQWKVDVGLTCDPNIYPAKTVPYLTPEEATELAFGAQVLAHQSMRPARBGDI 360
 QY 360 PVVKSYNPKAPGTLITRQDMDXGLVLTSLVLSNVTMLDIVSTRMLGOYGFARVSG 419
 DB 360 PVVKSYNPKAPGTLITRQDMDXGLVLTSLVLSNVTMLDIVSTRMLGOYGFARVSG 418
 QY 419 AII-PEDLCISVDCVATSEVSVSLSDSKISRELIQASSELDHVVELEKIAIVRLIQ 474
 DB 419 AII-PEDLCISVDCVATSEVSVSLSDSKISRELIQASSELDHVVELEKIAIVRLIQ 474
 QY 480 ORAIIISLIGNVEQSSLIIEKTGRVLRKSGVNVQMIISQASKVMNLSLIVHSDAKALVEAL 539
 DB 475 ORAIIISLIGNVEQSSLIIEKTGRVLRKSGVNVQMIISQASKVMNLSLIVHSDAKALVEAL 534
 QY 540 HOAFFEDDVLSCVEAENLLVG 560
 DB 535 HOAFFEDDVLSCVEAENLLVG 555

RESULT 3

AAE04359
 ID AAE04359 standard; protein; 564 AA.

XX

AC AAE04359;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Glycine max aspartate kinase.
 XX
 XX Soy bean; aspartate kinase; cell transformation; transgenic plant.
 XX
 OS Glycine max.
 XX
 PN W0200146393-A2.
 XX
 PD 28-JUN-2001.
 XX
 PF 19-DEC-2000; 2000MO-US034396.
 XX
 PR 21-DEC-1999; 99US-0172944P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Falco SC, Farnodu CO, Thorpe CJ;
 XX
 DR WPI; 2001-418057/44.
 XX
 PT Novel nucleic acid molecules encoding aspartate kinase useful in cell
 XX transformation and transgenic plant production procedures.
 XX
 PS Example 3; Fig 1; 64pp; English.
 XX
 CC The present invention relates to nucleic acid molecules encoding
 CC aspartate kinase useful in cell transformation and transgenic plant
 CC production procedures. Nucleic acid encoding aspartate kinase is useful
 CC for producing a transgenic plant, by transforming a plant cell with a
 CC chimeric gene comprising nucleic acid encoding aspartate kinase operably
 CC linked to a regulatory sequence and regenerating a plant from the
 CC transformed plant cell. The present amino acid sequence is Glycine max
 CC aspartate kinase
 CC
 SQ Sequence 564 AA;
 XX
 Query Match 65.2%; Score 1835; DB 4; Length 564;
 Best Local Similarity 70.6%; Pred. No. 2.3e-164;
 Matches 389; Conservative 42; Mismatches 80; Indels 40; Gaps 7;
 37 GPRGARGLSWVADSTSRRAKQADGGVLAGAPVL----- 71
 10 GVGQKLAVSWSVRSILHCKSQI--GFAALGAPVCARVWGNRAVSVTTCKASTDYIE 67
 72 -----GGLMEGLGDLQSVYMTFGSSSVSAAARAEVGLITPPERPVYVLSAMGT 125
 68 KXATENGWVSSSG-ETSFCTVMKFGSSVASADRMKEVATLLSPPEEPVIVLSAMGT 126
 126 TNNLLAGEKAVGCGVIVHSEIEEMMVVSLHKTVDGLP-XICNTSLYELQELKGI 184
 127 TNNLLAGEKAVGCGVIVHSEIEELPDIHRTVDQGVGVSISKLEELQELKGI 186
 185 AMMKELTPRTSDIVSGECMSTRISATLANKRVRARROYDAFDIGITTTDFEGNADILE 244
 187 AMMKELTKRQDILVSGECMSTRIFAALINKGVARQYDAFEIGFITTDFTADILE 246
 245 ATTPAVAKRLHGMIDPAIPVVTGFLGKWKSGAVTTIGRGSDLTATTICKALGLEI 304
 247 ATTPAVAKRLHGMISPAIATVTFGLKARKSCAVTTIGRGSDLTATTICKALGLEI 306
 305 QVWADVUGVITCDPNITPAKTVPYLTPEEATLAFGAQVTLHPQSMRPARAGDIPVAVK 364
 307 QVWADVUGVITCDPNITPAKTVPYLTPEEATLAFGAQVTLHPQSMRPARAGDIPVAVK 366
 365 NSYPAKPGTLITQRDMXGLVLTSLVSKSVNTMLDVSTRMLGQYGLFARVSGICVY 424
 367 NSYPAKPGTLITQARMK--AVLTSIVLKRVTMLDLASTRMLCQYFLAKVSI--F 422
 425 EDLCISVDCAVTSEVSVSLDPSKIMRELIIQASELDHVVLEELKAIYVLLQORAI 484

DB 423 EELGISVDVATSEVSVSLTDPDKLMSRELIIQASELDHVVLEELKAIYVLLQORAI 482
 QY 485 SLIGNVEOSSLILEKTGTVLRKSGVNVGMISGASKVNMSLIYHDSAKALVEALHQAFF 544
 DB 483 SLIGNVQSSLILEKTLVLRKSGVNVGMISGASKVNMSLIYHDSAKALVEALHQAFF 542
 QY 545 EDDVLSQVDAE 555
 DB 543 ESE-LSELEMD 552
 RESULT 4
 ID ABB93449
 ID ABB93449 standard; protein; 569 AA.
 AC ABB93449;
 XX
 DT 31-MAY-2002 (first entry)
 XX
 DE Herbicidally active polypeptide SEQ ID NO 2660.
 XX
 KW Herbicidal; plant; agriculture; herbicide.
 XX
 OS Arabidopsis thaliana.
 XX
 PN W0200210210-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 28-AUG-2001; 2001WO-EP009892.
 XX
 PR 28-AUG-2001; 2001WO-EP009892.
 XX
 PA (PARB) BAYER AG.
 XX
 PI Tietjen K, Weidler M;
 XX
 DR WPI; 2002-269010/31.
 XX
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.
 XX
 PS Claim 5; SEQ ID NO 2660; 261pp + Sequence Listing; English.
 XX
 CC The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides
 CC
 SQ Sequence 569 AA;
 XX
 Query Match 63.6%; Score 1789; DB 5; Length 569;
 Best Local Similarity 71.8%; Pred. No. 5.4e-160;
 Matches 369; Conservative 54; Mismatches 79; Indels 12; Gaps 6;
 44 LSWVADSTSRRAKQADGGVLAGAPVLGGLMEGLGD-----QLSVWKFQSSSVSAAAR 99
 44 LSLPIDGSSIRKVSQSSGRNIVRA-VLEKXTEAITVEDKGITCWKFGSSVVAAR 102
 100 MAEVAGLITPPEERPVVLSAMGKTNNLLAGEKAVGCGVIVHSEIEEMNVKSLHIK 159
 103 MKEVADLITPPEESPVVLSAMGKTNNLLAGEKAVGCGVIVHSEIEELSIKEDHIR 162
 160 TVDELGL-XICNTSLYELQELKGIAMMKELTPRTSDIVSGECMSTRISATLANKR 218

Accession	Query Match	Score	Length	DB	Similarity	Conservative	Mismatches	Indels	Gaps
PR 25-OCT-1999;	99US-0161404P.	74.5%	544	DB 3	62.1%	50	65	5	3
PR 25-OCT-1999;	99US-0161405P.								
PR 25-OCT-1999;	99US-0161406P.								
PR 25-OCT-1999;	99US-0161359P.								
PR 25-OCT-1999;	99US-0161360P.								
PR 25-OCT-1999;	99US-0161361P.								
PR 25-OCT-1999;	99US-0161362P.								
PR 25-OCT-1999;	99US-0161363P.								
PR 25-OCT-1999;	99US-0161364P.								
PR 25-OCT-1999;	99US-0161365P.								
PR 25-OCT-1999;	99US-0161366P.								
PR 25-OCT-1999;	99US-0161367P.								
PR 25-OCT-1999;	99US-0161368P.								
PR 25-OCT-1999;	99US-0161369P.								
PR 25-OCT-1999;	99US-0161370P.								
PR 25-OCT-1999;	99US-0161371P.								
PR 25-OCT-1999;	99US-0161372P.								
PR 25-OCT-1999;	99US-0161373P.								
PR 25-OCT-1999;	99US-0161374P.								
PR 25-OCT-1999;	99US-0161375P.								
PR 25-OCT-1999;	99US-0161376P.								
PR 25-OCT-1999;	99US-0161377P.								
PR 25-OCT-1999;	99US-0161378P.								
PR 25-OCT-1999;	99US-0161379P.								
PR 25-OCT-1999;	99US-0161380P.								
PR 25-OCT-1999;	99US-0161381P.								
PR 25-OCT-1999;	99US-0161382P.								
PR 25-OCT-1999;	99US-0161383P.								
PR 25-OCT-1999;	99US-0161384P.								
PR 25-OCT-1999;	99US-0161385P.								
PR 25-OCT-1999;	99US-0161386P.								
PR 25-OCT-1999;	99US-0161387P.								
PR 25-OCT-1999;	99US-0161388P.								
PR 25-OCT-1999;	99US-0161389P.								
PR 25-OCT-1999;	99US-0161390P.								
PR 25-OCT-1999;	99US-0161391P.								
PR 25-OCT-1999;	99US-0161392P.								
PR 25-OCT-1999;	99US-0161393P.								
PR 25-OCT-1999;	99US-0161394P.								
PR 25-OCT-1999;	99US-0161395P.								
PR 25-OCT-1999;	99US-0161396P.								
PR 25-OCT-1999;	99US-0161397P.								
PR 25-OCT-1999;	99US-0161398P.								
PR 25-OCT-1999;	99US-0161399P.								
PR 25-OCT-1999;	99US-0161400P.								
PR 25-OCT-1999;	99US-0161401P.								
PR 25-OCT-1999;	99US-0161402P.								
PR 25-OCT-1999;	99US-0161403P.								
PR 25-OCT-1999;	99US-0161404P.								
PR 25-OCT-1999;	99US-0161405P.								
PR 25-OCT-1999;	99US-0161406P.								
PR 25-OCT-1999;	99US-0161407P.								
PR 25-OCT-1999;	99US-0161408P.								
PR 25-OCT-1999;	99US-0161409P.								
PR 25-OCT-1999;	99US-0161410P.								

XX	WP1; 2002-269010/31.
XX	
PT	Identifying plant target proteins for herbicidally active compounds.
PT	Computing aligning and comparing nucleic acid or amino acid sequences
PT	from plant with nucleic acid or amino acid sequences from non-plant
PT	organisms.
PS	Claim 5; SEQ ID NO 2671; 26tp + Sequence Listing; English.
XX	
CC	The invention relates to identifying target proteins (ABB90790-ABB94016)
CC	for herbicidally active compounds, comprising aligning and comparing
CC	nucleic acid or amino acid sequences from plant with nucleic acid or
CC	amino acid sequences from non-plant organisms using suitable search
CC	parameters, where plant sequences having an E-value greater by a factor
CC	of 3 than the E-value of most similar non-plant sequences are selected.
CC	The polypeptides or nucleic acids encoding them are useful for
CC	identifying modulators. The identified modulators are useful as
CC	herbicides
SQ	Sequence 544 AA;
Query Match	62.1%; Score 1747.5; DB 5; Length 544;
Best Local Similarity	74.5%; Pred. No. 4.3e-156;
Matches 351; Conservative 50; Mismatches 65; Indels 5; Gaps 3	
Db	78 GLGDOLSYVMKFGGSSVSSAARMAEAVAGILTFPEERPVYVLAMGKTNNLLAAEKAV 137
OY	138 GCGVIVHSEIEEMNMWYSLHIKTVDLGL-PXICNTSLYELEQLKGIAWKELPRISD 196
Db	137 TCGVTWVESIEESLFKEHLRTAEHELGETTVLEKHLEGHOLLKGISMKKELTERTD 196
OY	197 YLVSPFGCWSTRFSFYLTKIKVRKROYDAPIGITTFDEGNADILEATYPVAARLHG 256
Db	197 YLVSPFGCWSTRFSFYLTKIKGHKAROYDAPEIGITTTDDPYTNADILEATYPVASKTLWG 256
OY	257 DWIQDPALPVVTGFLICKGMKSGAVTTLRGGSDLTATTIGKALGREIQWCKDVGYLTLC 316
Db	257 DWSKENAPVPVATGYLKRGARSQAITTLTGGSGLDTATTIGALGREIQWCKDVGYLTLC 316
OY	317 DPNIYPAKVPLPTEFEATELAYGAOVLHPQSMPAREBDIVRKNSYNPKAGSTLI 376
Db	317 DPNIYPAQGVPLUTIDEALAYGAOVLHLSMRPARDDIVRKNSYNPTAGTYI 376
OY	377 TRQRMDXGLVLTSLVLSKNVTMLDIVSTRMLGOYGFLARVSICYIEDLCISVDCVAT 436
Db	377 TRSRDMGR-AVLTSLVLRNVTMTDLASTRLMGYGFGLAV--FTTFEDLGISVVVAT 432
OY	437 SEVSASVSLDPKIMSRELIIQOASELDHYVELEKIATVRLIOORAIISLIGNEOSSLI 496
Db	433 SEVISITLTPALMGRRELIQRNELNDNVETLEKIAVVKLQGRSIIISLIGNVOSSLI 492
OY	497 LEKTRGVLRKSUNNVOMISOGASKRMCSLIYDSPAKLVEALHQAFEDD 547
Db	493 LKRVFYFRSNGVNVMISOGASKVMISLIVNDEAEQCVALLSAFFEDD 543
RESULT 7	
AAG46215	
ID	AAG46215 standard; protein; 563 AA.
XX	
AC	AAG46215;
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 58120.
XX	
KM	Protein identification; signal transduction pathway; metabolic pathway;
KM	hybridisation assay; genetic mapping; gene expression control; promoter;
XX	termination sequence.
XX	

OS	Arabidopsis thaliana.	
XX	EP1033405-A2.	
XX	06-SEP-2000.	
XX	25-FEB-2000;	2000EB-00301439.
XX	25-FEB-1999;	99US-0121825P.
XX	05-MAR-1999;	99US-0123180C.
XX	09-MAR-1999;	99US-0123548B.
XX	23-MAR-1999;	99US-0125588P.
XX	25-MAR-1999;	99US-0126264P.
XX	29-MAR-1999;	99US-0126785P.
XX	01-APR-1999;	99US-0127626P.
XX	06-APR-1999;	99US-0128934P.
XX	08-APR-1999;	99US-0128714P.
XX	16-APR-1999;	99US-0129074P.
XX	19-APR-1999;	99US-0130077P.
XX	21-APR-1999;	99US-0130449P.
XX	23-APR-1999;	99US-0130510P.
XX	23-APR-1999;	99US-0130891P.
XX	28-APR-1999;	99US-0131449P.
XX	30-APR-1999;	99US-0132048P.
XX	30-APR-1999;	99US-0132407P.
XX	04-MAY-1999;	99US-0132484P.
XX	05-MAY-1999;	99US-0132885P.
XX	06-MAY-1999;	99US-0132486P.
XX	07-MAY-1999;	99US-0132487P.
XX	11-MAY-1999;	99US-0132863P.
XX	14-MAY-1999;	99US-0134356P.
XX	14-MAY-1999;	99US-0134318P.
XX	14-MAY-1999;	99US-0134411P.
XX	14-MAY-1999;	99US-0134372P.
XX	14-MAY-1999;	99US-0134370P.
XX	18-MAY-1999;	99US-0134768P.
XX	19-MAY-1999;	99US-0134341P.
XX	20-MAY-1999;	99US-0135124P.
XX	21-MAY-1999;	99US-0135535P.
XX	25-MAY-1999;	99US-0135629P.
XX	27-MAY-1999;	99US-01363021P.
XX	28-MAY-1999;	99US-01365929P.
XX	01-JUN-1999;	99US-0137282P.
XX	03-JUN-1999;	99US-0137528P.
XX	04-JUN-1999;	99US-0137502P.
XX	07-JUN-1999;	99US-01373724P.
XX	08-JUN-1999;	99US-01386540P.
XX	10-JUN-1999;	99US-01388440P.
XX	14-JUN-1999;	99US-01386847P.
XX	16-JUN-1999;	99US-01391119P.
XX	17-JUN-1999;	99US-01394952P.
XX	17-JUN-1999;	99US-01394543P.
XX	18-JUN-1999;	99US-01394542P.
XX	18-JUN-1999;	99US-01394545P.
XX	18-JUN-1999;	99US-01394656P.
XX	18-JUN-1999;	99US-01394657P.
XX	18-JUN-1999;	99US-0139463P.
XX	18-JUN-1999;	99US-0139763P.
XX	18-JUN-1999;	99US-01398177P.
XX	21-JUN-1999;	99US-0139899P.
XX	23-JUN-1999;	99US-0140353P.
XX	23-JUN-1999;	99US-0140354P.
XX	24-JUN-1999;	99US-0140659P.
XX	29-JUN-1999;	99US-0140991P.

Db 435 V--FTTFEDLGISVDVATSEVSIITLIDPAKMGRELIIQRVNEIDNIVEELEKIAVVKL 492
Qy 478 LQQRALISLIGNVEOSLLEKTCGRVLRSGVNNMISOGASKYMSIIVHDSDAKLVE 537
Db 493 LQRRSIISLIGNQKSSLLEKVFQVFRSNGVNMISOGASKYMSIIVHDEABQCR 552
Qy 538 ALHQAFEDD 547
Db 553 ALHSAFFETD 562
RESULT 8
AAG46237
ID AAG46237 standard; protein; 459 AA.
XX AAG46237;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58149.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN BP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 14-MAY-1999; 99US-0134768P.
PR 16-MAY-1999; 99US-0134941P.
PR 19-MAY-1999; 99US-0135124P.
PR 20-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135629P.
PR 24-MAY-1999; 99US-0136021P.
PR 25-MAY-1999; 99US-0136392P.
PR 27-MAY-1999; 99US-0136782P.
PR 28-MAY-1999; 99US-0137222P.
PR 01-JUN-1999; 99US-0137528P.
PR 03-JUN-1999; 99US-0137582P.
PR 04-JUN-1999; 99US-0137724P.
PR 07-JUN-1999; 99US-0138094P.
PR 08-JUN-1999;

PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139465P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 19-JUL-1999; 99US-0144336P.
PR 19-JUL-1999; 99US-0144337P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145216P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

```

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148655P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 22-AUG-1999; 99US-0149902P.
PR 22-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151308P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157653P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159339P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 16-OCT-1999; 99US-0159684P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160982P.
PR 23-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

```

```

Query Match 61.4%; Score 1727.5; DB 3; Length 459;
Best Local Similarity 75.1%; Pred. No. 2.6e-154;
Matches 347; Conservative 48; Mismatches 62; Indels 5; Gaps 3;

```

```

Db 1 MKFGSSSVSARMAEVAAGLITFPERPVVIVLSANGKTTNILLAGEKAVGCGVTHSE 60
147 IEEEMNVXSLHIKIVDEIGL- PKICNTSLYEBOLKIGIAMKELTPTSDIYVSFECM 205
61 IEELSFIEKELHRIHAEIGVETVEIEKLEHQLLIGISMKELTTRIDYLVSFECM 120
206 STRIFSAYLNKIRVARQYDAFDIGFTTDEFGNADILEATPAVAXRLHGDWICDPAIP 265
121 STRLSAYLNKIGHARQYDAFEIGFTTDDFTNADILEATPAVAXKTIWGDMSKENV 180
Qy 266 VVTGFLGKMGSGAVTTIGRGSDITATTICKALGLRBIQWQVQVQVGLTCDPNIYHAK 325
Db 181 VVTGYLGKMGRCALITLGRGSDITATTICKALGLRBIQWQVQVQVGLTCDPNIYHQAQ 240
Qy 326 TVPYLTFEATLAFYQVYHPQSMRAREGDI PVRYKNSYNKAPGTLITRRDMDXG 385
Db 241 SVPYLTFEATLAFYQVYHPQSMRAREGDI PVRYKNSYNKAPGTLITRRDMDXK- 299
Qy 386 LVLTSLVKSNTVMTLDIVSTRMLGOYFLARVSGICYIEDLCISVDCVATSEVSVSL 445
Db 300 -AVLTSLVKSNTVMTLDIVSTRMLGOYFLAKV--FTTFEDLGISVDVATSEVSI 356
Qy 446 DEPKTWSRPLIQASBHDVVELEKIAVLLQORALISLIGVQOSLILEKTVLR 505
Db 357 DPAKLMGRFLIQRVNELNVELEKIAVLLQORALISLIGVQOSLILEKTVLR 416
Qy 506 KSGVNVQMSIGASKVNMSLIYHDSADALYEAALHQAFFEDD 547
Db 417 SNGVNVQMSIGASKVNISLIYHDEAAGCVAHSAFFEDD 458

RESULT 9
AAG46216
ID AAG46216 standard; protein; 478 AA.
XX
AC AAG46216;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58121.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EB-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 23-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131499P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.

```

PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0135494P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 01-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142302P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148313P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150889P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153738P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155466P.
PR 24-SEP-1999; 99US-0156659P.
PR 28-SEP-1999; 99US-0156465P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157783P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158223P.
PR 12-OCT-1999; 99US-0158363P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160747P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.

PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 29-OCT-1999; 99US-0162142P.

 Query Match 60.7%; Score 1708; DB 3; Length 478;
 Best Local Similarity 72.4%; Pred. No. 1,9e-152;
 Matches 347; Conservative 48; Mismatches 62; Indels 24; Gaps 4;

 QY 87 MKFGSSVSSAAMAEVAGILTFPERPVVLTLSAMGKTTNLLAGEKAVGCVIHTSE 146
 1 MKFGSSVSSAAMAEVAGILTFPERPVVLTLSAMGKTTNLLAGEKAVGCVIHTSE 60
 QY 147 IEEMNVKSLH-----KTVDLGL-FCICNTSLYELEQLIKGIAM 186
 61 IEELSFKEHLHVSHPNLALFSEWIKFSGMFRFAHLEGVETTVIEKHLEGLHQLIKGISM 120
 QY 187 MKELTPTSDYLVSFGECMSTRFFSAVLRKIRKARQYAFDIFGTTDEFGNADILEAT 246
 121 MKELTPTSDYLVSFGECMSTRFFSAVLRKIRKARQYAFDIFGTTDEFGNADILEAT 180
 QY 247 YPAVAKRLHGDWITODPALPVYTGFLGKWKSGAVTTLGRGSDLTATTIGKALGREIQV 306
 181 YPAVAKRLHGDWITODPALPVYTGFLGKWKSGAVTTLGRGSDLTATTIGKALGREIQV 240
 QY 307 WKQVVDGVLTCDDPNTIYHAKTVPLTFEETELAYFAQVLIHQSNRPARREGDIPVRVNS 366
 241 WKQVVDGVLTCDDPNTIYHAKTVPLTFEETELAYFAQVLIHQSNRPARREGDIPVRVNS 300
 QY 367 YNPAPGTLITRORDMDXGLVLTSLVLSKSNVMTLMDIVSTRMGQYFLARVSGICVYED 426
 301 YNPAPGTLITRORDMDXGLVLTSLVLSKSNVMTLMDIVSTRMGQYFLARVSGICVYED 356
 QY 427 LCISVDCVATSEVSVSDPSKIMSRRELIOQASGLDHYVELEKIALVRLIQORAIISL 486
 357 LGISVDCVATSEVSVSDPSKIMSRRELIOQASGLDHYVELEKIALVRLIQORAIISL 416
 QY 487 IGVNEOSVLTLEKTRVLRKSGVNVQMISQASKVMSLIYHSDAKALVEALHQAFFED 546
 417 IGVNEOSVLTLEKTRVLRKSGVNVQMISQASKVMSLIYHSDAKALVEALHQAFFED 476
 QY 547 D 547
 Db 477 D 477

 RESULT 10
 ID AAG46238 standard; protein; 446 AA.
 XX AAG46238;
 XX
 DT 18-OCT-2000 (first entry)
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 58150.
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KM hybridisation assay; genetic mapping; gene expression control; promoter;
 KM termination sequence.
 XX Arabidopsis thaliana.
 XX

PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 XX 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130445P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 04-MAY-1999; 99US-0132407P.
 PR 05-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135628P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 19-MAY-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142802P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0144624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144332P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144844P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147328P.
 PR 05-AUG-1999; 99US-014760P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148119P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148655P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-014926P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 25-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-015138P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.

PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155466P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156566P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158022P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159283P.
 PR 13-OCT-1999; 99US-0159284P.
 PR 13-OCT-1999; 99US-0159285P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 26-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 59.6%; Score 1675.5; DB 3; Length 446;
 Best Local Similarity 74.8%; Pred. No. 2.1e-149; Indels 5; Gaps 3;
 Matches 336; Conservative 48; Mismatches 60;

QY 100 MAEVAGLILTPPEERPVVLSAMGKTTNNLLAGEKAVGGVTHVSEIEEMNVKSLHKK 159
 DB 1 MKEVANLILSPDERPVVLSAMGKTNKLLKAGEKAVGCVNVSIEELSPKELHUR 60
 QY 160 TYDELGL-PRICNSLYVELBOLKIGIAMKELTPRTSDYVSEGECKSTRIFSAVYANKIR 218
 DB 160 TYDELGL-PRICNSLYVELBOLKIGIAMKELTPRTSDYVSEGECKSTRIFSAVYANKIR 218
 QY 61 TANELGVETTVIEKHEBGLHQLKIGISMKEELTRTDYVSEGECKSTRIFSAVYANKIR 120
 DB 61 TANELGVETTVIEKHEBGLHQLKIGISMKEELTRTDYVSEGECKSTRIFSAVYANKIR 120
 QY 219 VKARQYDAFDIGITTFTEFGNADILEATYPAVAKRLHGMWIODPAIPVVTGFLGKWKSG 278
 DB 121 HKARQYDAFEIGITTFDFTNADILEATYPAVSKTLVGDMKKNAPVVTGVLGKWRSC 180
 QY 279 ATTITGRGSDLTATTIKALGELIQWMDVGVLTCDNNTIPHKATVYLLFEEBATEL 338
 DB 181 ATTITGRGSDLTATTIKALGELIQWMDVGVLTCDNNTIPGASVYLLFDEBAEL 240
 QY 339 AFYGAQVLAPOSMPAREGDIPIVAKNSYNPKAPGTLITRORDMXGLVVLTSIVLKSXY 368
 DB 241 AFYGAQVLAPOSMPAREGDIPIVAKNSYNPKAPGTLITRORDMXGLVVLTSIVLKSXY 298
 QY 399 TMLDVTSTRMGQGFHLARVSGICIEDICISVUCVATSEVSUSVSLDPEKISRRELIQ 458
 DB 299 TMLDVTSTRMGQGFHLARVSGICIEDICISVUCVATSEVSUSVSLDPEKISRRELIQ 356
 QY 459 ASELDHVVLELEKIAIVLLOQRAIISLIGNEOSSILLEKGTGVLRKSGVNTOMISOGA 518
 DB 357 VNELDNVLELEKIAIVLLOQRAIISLIGNEOSSILLEKGTGVLRKSGVNTOMISOGA 416

QY 519 SKYMSLIVHSDAKALVEALHQAFFEDD 547
DB 417 SKYMSLIVHDEEACQCRALHSAFFETD 445

RESULT 11
ID AAG05075 standard; protein: 555 AA.
AAG05075;
17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 1342.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132464P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-014218P.
XX 14-MAY-1999; 99US-014219P.
XX 14-MAY-1999; 99US-014221P.
XX 14-MAY-1999; 99US-014370P.
XX 18-MAY-1999; 99US-0143768P.
XX 19-MAY-1999; 99US-0144941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137528P.
XX 01-JUN-1999; 99US-0137528P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139482P.

PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141872P.
PR 01-JUL-1999; 99US-0141872P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145513P.
PR 27-JUL-1999; 99US-0145818P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148584P.

PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151800P.
PR 30-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151388P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 59.64; Score 1675.5; DB 3; Length 555;
Best Local Similarity 73.54; Pred. No. 2.9e-149;
Matches 347; Conservative 47; Mismatches 71; Indels 7; Gaps 4;

QY 82 QLSVVMKFGGSSVSSAARMAEVAGLITPPEERPVVVISAMKTTNNILAGEKAVGCGV 141
DB 78 KLTCEVMKFGGSSVSSAARMAEVAQLITPPEERPVVVISAMKTTNNILAGEKAVGCGV 137
QY 142 IHVSEIEEMNVKSLHKTVDLGL-PIICNTSLVELBOLLGSIAMKELTPRTSDYLS 200
DB 138 TIVDTIELSLYIKELHIRTAEHGVETAVIAEHLBOLLGSIAMKELTPRTSDYLS 197

QY 201 FGEONSTIFSAAYLNKIRVAKROYDAPDIDGFIITTFDERGNADILEATYPVAKRLHGDWQ 260
DB 198 FGEONSTIRLFAAYLNKIGHKARQYDAFEIIGITTDFTNADILEATYPVAKSKLLJDDWSK 257
QY 261 DPAIPVMTGFLGKMGKGAVALTLGRGSDLTATTIGKALGREIQWKVDVGLTCDPNI 320
DB 256 ENALPVYTGFLGKMGKSCAVTLTLGRGSDLTATTIGKALGREIQWKVDVGLTCDPNI 317
QY 321 YEHAKTVPLTFEATELAFGAQVLIHPQSNRPAREGDIIVRYKSNYNPAFGTLITRQR 380
DB 318 YCGAQPVPHLTDEDAELAFGAQVLIHPLSRPAEENIIVRYKSNYNPAFGVITRQR 377
QY 381 DMDXGLVLTISIVKSNVTMLDIYSTMLGQGLARVSGICIEDICISVDCVATSEVS 440
DB 378 DMSK--AVLTISIVKSNVTMLDIYSTMLGQGLARV--FSTPEKLGISVDVATSEVS 433
QY 441 VVSVDPSKINSRELIQASSELGHVEELEKIALVRLIQAPATISLIGNVQSSLIEXT 500
DB 434 ISLTLDPSKCSRELIQH--ELDQVEELEKIAVNNILRRHSIISLIGNVQSSFILEKG 491
QY 501 GRVLRKSGVNTQMISOGASKVNSLIVHSDPAKALVEALQAFEDDVLISQV 552
DB 492 FVRLRTNGINQMISOGASKVNSLIVNDDAEHCVALHSAFETDTCEAV 543

RESULT 12
AAG05074
ID AAG05074 standard; protein; 559 AA.
XX
AC AAG05074;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1341.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
CS Arabidopsis thaliana.
OS
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123546P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129645P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130445P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135529P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143654P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144311P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145068P.
PR 21-JUL-1999; 99US-0145068P.
PR 22-JUL-1999; 99US-0145065P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146386P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 15-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149925P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150565P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151068P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152263P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154799P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157153P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.

PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 26-OCT-1999; 99US-0161362P.
 PR 28-OCT-1999; 99US-0161392P.
 PR 28-OCT-1999; 99US-0161393P.
 PR 28-OCT-1999; 99US-0162142P.

Query Match 59.6%; Score 1675.5; DB 3; Length 559;
 Best Local Similarity 73.5%; Pred. No. 2.9e-149;
 Matches 347; Conservative 47; Mismatches 71; Indels 7; Gaps 4;

QY 82 QLSVWKKFGSSVSSAARMAEVAAGLITFPEERPVVLSAMGKTTNNLLAGEKAVGCGV 141
 DB 82 KLTGVMKFGSSVSSAARMAEVAAGLITFPEERPVVLSAMGKTTNNLLAGEKAVGCGV 141
 QY 142 IHVSEIEEMNNVKSILHKTVDLGL-EXICNTSLYELEQLKGIAMKELTPRTSDYLV 200
 DB 142 TIVDTIEELSYKEHLIRTAHELGEVETAVIAHLEGLQKGVAMKELTPRSRDYLV 201
 QY 201 FEECWSTRIFSAVYLNKIRKARQYDAFDIGFITTDEFGNADILEATYPAVAKLHGDMIQ 260
 DB 202 FEECWSTRIFSAVYLNKIRKARQYDAFDIGFITTDEFGNADILEATYPAVAKLHGDMIQ 261
 QY 261 DPAIPVVTGFLGKMGKSGAVTTLGRGSDLTATTIGKALGREIQWKDVGVLTCDDPNI 320
 DB 262 ENALPVTGFLGKMGKSGAVTTLGRGSDLTATTIGKALGREIQWKDVGVLTCDDPNI 321
 QY 321 YPAKTVPLTFEEATELAFGAQVLPQSNRPAEEDIPVRVNSVNPAPGTLITROR 380
 DB 322 YCAQPVPLTFDEAEALAFGAQVLPQSNRPAEEDIPVRVNSVNPAPGTLITROR 381
 QY 381 DMDXGLVLTSLVSKSNVTMLDIVSTRMLGQYGFARVSGICYIEDLCISVDCVATSEVS 440
 DB 382 DMSK--AVLTSLVSKSNVTMLDIVSTRMLGQYGFARV--FSTFEKLGISVDVATSEVS 437
 QY 441 VSVSLDPSKIMSELIQOASELDHYVELEKIALVRLQORALISLIGNVQSSLLIEXT 500
 DB 438 ISLTLDPSKFCSEELIQH--ELDQVEELEKIAVNNLRRHSILSLIGNVQSSFLIEKG 495
 QY 501 GRVLRSGVNVQMISQASKVNMSLIVHSDAKALVEALHQAFFEDVLSQV 552
 DB 496 FRVLRNGINVMISQASKVNMSLIVHSDAKALVEALHQAFFEDVLSQV 547

RESULT 13

ABB92174
 ID ABB92174 standard; protein; 559 AA.

XX ABB92174;

XX 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 1385.

KM Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

XX W0200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-BP009892.

XX 28-AUG-2001; 2001WO-EP009892.

XX (FARB) BAYER AG.

XX Tietjen K, Weidner M;

XX

DR WPI, 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.

PS Claim 5; SEQ ID NO 1385; 261pp + Sequence Listing; English.

CC The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides

XX Sequence 559 AA;

Query Match 59.6%; Score 1675.5; DB 5; Length 559;
 Best Local Similarity 73.5%; Pred. No. 2.9e-149;
 Matches 347; Conservative 47; Mismatches 71; Indels 7; Gaps 4;

QY 82 QLSVWKKFGSSVSSAARMAEVAAGLITFPEERPVVLSAMGKTTNNLLAGEKAVGCGV 141
 DB 82 KLTGVMKFGSSVSSAARMAEVAAGLITFPEERPVVLSAMGKTTNNLLAGEKAVGCGV 141
 QY 142 IHVSEIEEMNNVKSILHKTVDLGL-EXICNTSLYELEQLKGIAMKELTPRTSDYLV 200
 DB 142 TIVDTIEELSYKEHLIRTAHELGEVETAVIAHLEGLQKGVAMKELTPRSRDYLV 201
 QY 201 FEECWSTRIFSAVYLNKIRKARQYDAFDIGFITTDEFGNADILEATYPAVAKLHGDMIQ 260
 DB 202 FEECWSTRIFSAVYLNKIRKARQYDAFDIGFITTDEFGNADILEATYPAVAKLHGDMIQ 261
 QY 261 DPAIPVVTGFLGKMGKSGAVTTLGRGSDLTATTIGKALGREIQWKDVGVLTCDDPNI 320
 DB 262 ENALPVTGFLGKMGKSGAVTTLGRGSDLTATTIGKALGREIQWKDVGVLTCDDPNI 321
 QY 321 YPAKTVPLTFEEATELAFGAQVLPQSNRPAEEDIPVRVNSVNPAPGTLITROR 380
 DB 322 YCAQPVPLTFDEAEALAFGAQVLPQSNRPAEEDIPVRVNSVNPAPGTLITROR 381
 QY 381 DMDXGLVLTSLVSKSNVTMLDIVSTRMLGQYGFARVSGICYIEDLCISVDCVATSEVS 440
 DB 382 DMSK--AVLTSLVSKSNVTMLDIVSTRMLGQYGFARV--FSTFEKLGISVDVATSEVS 437
 QY 441 VSVSLDPSKIMSELIQOASELDHYVELEKIALVRLQORALISLIGNVQSSLLIEXT 500
 DB 438 ISLTLDPSKFCSEELIQH--ELDQVEELEKIAVNNLRRHSILSLIGNVQSSFLIEKG 495
 QY 501 GRVLRSGVNVQMISQASKVNMSLIVHSDAKALVEALHQAFFEDVLSQV 552
 DB 496 FRVLRNGINVMISQASKVNMSLIVHSDAKALVEALHQAFFEDVLSQV 547

RESULT 14

AAG05076
 ID AAG05076 standard; protein; 473 AA.

XX AAG05076;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 1343.

KM Protein identification; signal transduction pathway; metabolic pathway;

KM hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

XX Arabidopsis thaliana.

OS

XX
PN BP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134421P.
PR 14-MAY-1999; 99US-0134421P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135699P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136382P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137582P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140685P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142056P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0143977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145214P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147535P.
PR 10-AUG-1999; 99US-0148171P.
PR 10-AUG-1999; 99US-0148319P.
PR 11-AUG-1999; 99US-0148341P.
PR 12-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 18-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0152330P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.

```

PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156586P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158393P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 23-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161599P.
PR 26-OCT-1999; 99US-0161606P.
PR 26-OCT-1999; 99US-0161611P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

```

Query Match 59.2%; Score 1666.5; DB 3; Length 473;
 Best Local Similarity 73.9%; Pred. No. 1.6e-148;
 Matches 345; Conservative 45; Mismatches 70; Indels 7; Gaps 4;

```

QY 87 MKFGSSVSSAAMAAVAGILTFPERPVVVSAMGXTNNLLAGEKAVGCVIHYSE 146
DB 1 MKFGSSVSAAEKVIQVAKILTFPDEKPVVVSAMAKTNNLLMAGEKAVCGVTNVDI 60
QY 147 IEEWNVKSLHIKTVDLGL-PXICNTSLYEISQLIKGIAMKELTPTSDYLVFGEQM 205
DB 61 IEEISYIKELHIRTAHLEIVETAVIAEHLEGLQLKGVAMKELTLRSRDYLVFGEQM 120
QY 206 STRFSAVYLNKIRKARQVADAFIGFTTDFEGNADILEATYPAVAKRLHGMIDPAIP 265
DB 121 STRLFAVYLNKIGKARQVDAFELGITTDDFNAILLENTTPANKSKLIDMSKGNALP 180
QY 266 VVTGFLGKWKSGAVTTLGRGSDLTATTIGKALGREIQWKVDVGVLTCDPNTYPAK 325
DB 181 VVTGFLGKWRSCAVTTLGRGSDLTATTIGKALGREIQWKVDVGVLTCDPNTYCGAQ 240
QY 326 TVPYLTPEEATELAYPAQVTLHPOSMAPARESDIPRVKNSVYPKAPGTLITRODMDXG 385
DB 241 PVPHLTPEEAELAYPAQVTLHPSMPARESGIPRVKNSVYTPAPGVITRSRDMGR- 299
QY 386 LVVLTSLVKSNTMDIVSTRMLGQYGLFARVSGICYLEDCISVDCVATSEVSVSL 445
DB 300 -AVLTSLVKNVMTLDTSTRMLGQYGLFARV--FSTPEKLGISVDVATSEVSIITL 356
QY 446 DPKTMSRELIQASSELDHVVEELKTAIVRLIQARLISLGNVQSSLIIEKRGVVR 505
DB 357 DPKFCSRELIQH--ELDQVEELKTAIVNLRHSIISLIGNVQSSFIIEKGRVVR 414

```

```

QY 506 KSGVNVQMSQASKVMKSLIVHSDAKALVEALHQAFFEDPYLSQY 552
DB 415 TNGINVQMSQASKVMKSLIVNDDEAHECVKALHSAFFETDCEAY 461

RESULT 15
AA646217
ID AA646217 standard; protein; 465 AA.
AC
XX AA646217;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58122.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0122548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0128645P.
XX 18-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138647P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.

```



```
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141442P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142303P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143442P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 19-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 20-JUL-1999; 99US-0144614P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145087P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 22-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 26-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 27-JUL-1999; 99US-0145951P.  
PR 28-JUL-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147203P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147320P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 09-AUG-1999; 99US-0147936P.  
PR 10-AUG-1999; 99US-0148311P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148365P.  
  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149178P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150565P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151938P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153785P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157533P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159325P.  
PR 14-OCT-1999; 99US-0159328P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 18-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 26-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
  
Query Match 58.9%; Score 1656; DB 3; Length 465;  
Best Local Similarity 71.8%; Pred. No. 1,5e-147;  
Matches 336; Conservative 48; Mismatches 60; Indels 24; Gaps 4;  
  
Qy 100 MAEVAGLITPPEBRRVVLSSMGKTNNLLAGKAYCCGVHSELEENMKKSHI-158  
Db 1 MKEVANLISFDERVIVLSAMGKTNNLLAGKAYCCGVHSELEENMKKSHI-60  
Oy 159 -----KTVDELGL-PTICNTSLVELEOLIKGIAMKKELTPTSDYLV-199
```

Db 61 VHSPLALFSEWEIKFSMPTAHGLQVETTVIEKLEGLHQLKGISMMKELLTRDYLV 120
QY 200 SFGECMSTRIFSAVYANKIRVYAROYDAFDIGFITTDEFGNADILEATYPVAKRHHQDWI 259
121 SFGECMSTRIFSAVYANKIRVYAROYDAFDIGFITTDEFGNADILEATYPVAKRHHQDWI 180
QY 260 QDPALPVVYGLGKMGKSGAVTTLGRGSDLTATTIGKALGREIOVMKDVGVLTCDPN 319
181 KENAVPVVYGLGKMGKSGAVTTLGRGSDLTATTIGKALGREIOVMKDVGVLTCDPN 240
QY 320 IYPAKTVPYLTFEEATELAFYGAQVLEPQSMRPARBGDIPIRYKNSYNPAPGTLITRQ 379
241 IYPAQSVYPLTFDEPAELAYGAQVLEPQSMRPARBGDIPIRYKNSYNPAPGTLITRQ 300
QY 380 RMDXGLVYLTSLVLSKNTVMDLIVSTRMLGQYGFARVSGICYIEDLCISVDCVATSEV 439
301 RDMSK--AVLTSLVLSKNTVMDLIVSTRMLGQYGFARVSGICYIEDLCISVDCVATSEV 356
QY 440 SVSVSLDPKXINSRELIQASLELDHVELEKIALVRLQORATISLIQNTVESSLILEK 499
357 SISLTLDPAKMGRELQVNELDNLVELEKIAVAVKLLQRRSIIISLIQNVQKSSLILEK 416
QY 500 TGSVLRKSGVNTQMSIQGASKYNMSLIYHSDPAKALVEALHQAFFEDD 547
Db 417 VQVFRSNGVNTQMSIQGASKYNMSLIYHSDPAKALVEALHQAFFEDD 464

Search completed: March 23, 2004, 13:15:47
Job time : 85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:14:31 ; Search time 24 Seconds
(without alignments)

1204.606 Million cell updates/sec

Title: US-09-890-813-6

Sequence: 1 MAIPVRSAAAPRLVPSIPP.....QAFEDVLSQVEANLLVG 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6E_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6E_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	590	21.0	453	US-09-489-039A-8611	Sequence 8611, Ap
2	569	20.2	449	US-08-256-136-2	Sequence 2, Appl
3	569	20.2	449	US-08-950-737-2	Sequence 2, Appl
4	569	20.2	449	US-08-973-461A-8	Sequence 8, Appl
5	569	20.2	449	US-08-648-010-8	Sequence 8, Appl
6	565	20.1	460	US-09-543-681A-7248	Sequence 7248, Ap
7	469.5	16.7	835	US-09-489-039A-8740	Sequence 8740, Ap
8	469.5	16.5	820	US-08-380-182-23	Sequence 23, Appl
9	458.5	16.2	862	US-09-543-681A-6315	Sequence 6315, Ap
10	458.5	15.2	478	US-09-107-532A-6030	Sequence 6030, Ap
11	416.5	14.8	863	US-08-380-182-19	Sequence 19, Appl
12	416.5	14.8	463	US-09-134-001C-4160	Sequence 4160, Ap
13	414.5	14.7	421	US-08-532-828B-3	Sequence 3, Appl
14	414.5	14.7	421	US-08-532-828B-4	Sequence 4, Appl
15	414.5	14.7	421	US-08-700-359-9	Sequence 9, Appl
16	414.5	14.7	421	US-08-596-366-6	Sequence 6, Appl
17	414.5	14.7	421	US-08-967-104-6	Sequence 6, Appl
18	414.5	14.7	421	US-08-985-908-5	Sequence 5, Appl
19	414.5	14.7	421	US-08-853-730-14	Sequence 14, Appl
20	414.5	14.7	421	US-08-985-916-5	Sequence 5, Appl
21	414.5	14.7	863	US-08-380-182-20	Sequence 20, Appl
22	402	14.3	411	US-07-684-135A-2	Sequence 2, Appl
23	402	14.3	421	US-08-311-731A-30	Sequence 30, Appl
24	377.5	13.4	461	US-09-252-991A-29720	Sequence 29720, A
25	366.5	13.0	438	US-09-328-352-6511	Sequence 6511, Ap
26	366.5	13.0	828	US-09-489-039A-12219	Sequence 12219, A
27	334.5	11.9	441	US-09-198-452A-1124	Sequence 1124, Ap

28	325	11.6	424	US-09-134-001C-5225	Sequence 5225, Ap
29	300	10.7	815	US-09-543-681A-8330	Sequence 8330, Ap
30	294.5	10.5	285	US-09-540-236-3031	Sequence 3031, Ap
31	246	8.7	256	US-09-134-000C-6771	Sequence 6771, Ap
32	196	7.0	262	US-08-887-534A-23	Sequence 23, Appl
33	196	7.0	262	US-09-527-431-23	Sequence 23, Appl
34	187.5	6.7	156	US-09-056-556-227	Sequence 227, App
35	187.5	6.7	156	US-09-072-967-227	Sequence 227, App
36	187.5	6.7	156	US-09-072-967-227	Sequence 227, App
37	116	4.1	528	US-09-328-352-6385	Sequence 6385, Ap
38	108.5	3.9	250	US-09-543-681A-6626	Sequence 6626, Ap
39	105.5	3.8	432	US-09-252-991A-18147	Sequence 18147, A
40	105	3.7	1705	US-08-669-785-4	Sequence 4, Appl
41	104.5	3.7	702	US-09-328-352-8176	Sequence 8176, Ap
42	104	3.7	469	US-08-985-335-5	Sequence 5, Appl
43	102.5	3.6	528	US-09-410-372-5	Sequence 5, Appl
44	102.5	3.6	528	US-09-134-001C-4262	Sequence 4262, Ap
45	102.5	3.6	582	US-09-252-991A-25366	Sequence 25366, A

ALIGNMENTS

RESULT 1
US-09-489-039A-8611
Sequence 8611, Application US/09489039A
Parent No. 6610816

GENERAL INFORMATION:
APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 8611
LENGTH: 453
TYPE: PRT
ORGANISM: Klebsiella pneumoniae

US-09-489-039A-8611

Query Match 21.0%; Score 590; DB 4; Length 453;
Best Local Similarity 34.9%; Pred. No. 9, 2e-51;

Matches 166; Conservative 81; Mismatches 184; Indels 44; Gaps 15;

QY	85	VVMKFGGSVSAAARMAEVAAGLITFPERPVPVLSAMGKTNNLL-LAGEKAVCGVTH	143
DB	9	VMAKFGGTSVADPDANRSDIVALLDANTR-IVLSASAGVNIIVALLAG-----GLEP	61
QY	144	VSEIEENNVKSHIKTDELGLPXICNTSLVEBOLKGMKELPRTS-----DYL	198
DB	62	TERSQDALRQIOFNILRLRYPVIRE---EIERLINITTLAEAMALASTALTDEL	118
QY	139	VSPEECSTIFSAVINKIRKARQYDARDIGTITDEGND-----LLEATPAVAKR	253
DB	119	VSHLELSTLFEILRERGIQWDPDAKV-LRTNDRGRAPDPAIAAELTQQLAR	177
QY	254	LHGWDIDPAIPVVTGFLGKMGSGAVTTLGRGSDLTATTGKALGREIVMKVDGV	313
DB	178	L-----AEGVVTGIGISEAK-GRITTLGRGSDYTAALGALNATRVIMVDVPI	230
QY	314	LTCPPNYPKAKTVPIYTFEEATLVEFGAOVHPSQMPAREGDIPIVVKSYNPKAF	373
DB	231	YTPDPAAPAKRIDVIAFEAAEMATFGAKVHPATLPAVRSIDIPVVGSKPKAG	290
QY	374	TLITQRDMXGLVLTSLVLSKSNVTMLDIVSTRMLGQVGFARVSGICYBLCISYDC	433
DB	291	TLVKTKE---NPLFPALALRRQTLTLTSLNMLHSGFLAEVGGILARHN--ISVDL	345
QY	434	VATSEVSVSLDPSKTKNR--ELIQASLHVVELEKIAIVLQORALISLIGN-V	490

Db 346 TTTSVSVALLTMDTGSTSAGDTLLTQA-----LITELSSICRVEVEENLALVALIGNEL 400

QY 491 ESSLILEKTRGVLRKSGVNVQMISQASKYNMGLIYHSDPAKLYEALHQAFFE 545

Db 401 SNAQGVKEVFGVLEP--FNIRMLCYGASHNLGFLVPGDAEKVQKLNHNLF 453

RESULT 2

US-08-256-136-2

/ Sequence 2, Application US/08256136

/ Patent No. 5661012

/ GENERAL INFORMATION:

/ APPLICANT: SANO, KONOSUKE

/ APPLICANT: KOJIMA, HIROYUKI

/ APPLICANT: OGAMA, YURI

/ TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF L-THREONINE

/ NUMBER OF SEQUENCES: 4

/ CORRESPONDENCE ADDRESSES:

/ ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT

/ STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

/ CITY: ARLINGTON

/ STATE: VIRGINIA

/ COUNTRY: USA

/ ZIP: 22202

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patent In Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/256,136

/ FILING DATE: 01-JUL-1994

/ CLASSIFICATION: 435

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: PCT/JP93/01640

/ FILING DATE: 10-NOV-1993

/ PRIOR APPLICATION DATA: JP 300021/92

/ FILING DATE: 10-NOV-1992

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (703) 413-3000

/ TELEFAX: (703) 413-2220

/ TELEX: 248955 OPAT UR

/ INFORMATION FOR SEQ. ID NO: 2:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 449 amino acids

/ TYPE: amino acid

/ TOPOLOGY: linear

/ MOLECULE TYPE: protein

US-08-256-136-2

Query Match 20.2%; Score 569; DB 1; Length 449;

Best Local Similarity 34.0%; Pred. No. 1.2e-48;

Matches 161; Conservative 89; Mismatches 182; Indels 42; Gaps 14;

QY 85 VVWKFGSSSVSAARAAVAGLITPPEERPVVLSAMGKTNNLLAGEKAVCGVIHV 144

Db 5 VVSKFGSTVADPDANRSADIVLSDANVR-LVLSASAGITNLVALAE-----GLEPG 58

QY 145 SEIEEMNWKSLHIKIVDELGLPKICNTSLYEQLKGIAMKE-----LTPRTSDYLV 199

Db 59 EREKLDARNIQFALLERLRYP--NVIREIERLENTIVLAALALATSPALTDLV 115

QY 200 SPEECMSTRIFSAYLKRIRKARQYADPDIQITTFDEFGNAD-----LLEATYPAVAKRL 254

Db 116 SHELMSTLLFVEILNERDVQAWFVRKV-WRTNDRFGAPDPAALAEALQCLPRL 174

QY 255 HGMWIDPALPVYTGILGKWKSGAVTTIGRGSGLIKATTIKALGLEIOWMDVQVTL 314

Db 175 -----NEGIVITQFISGENK-GRITTLGRGGSYTAALLAEALHASRVIMTDVPGIY 227

QY 315 TCDPNIIYPAKTVPIYLTPEATELAFGAQVHLQPMRPARSGDIPVAVKSNYPKAPGT 374

Db 228 TTDPRIYSAKRIDEIAFAEAEMATFGAKVLPATLLPVAVSDIPVFGVSKSDPRAGT 287

QY 375 LITQGRMDKGLVLTSTIVLAKSVNMDIVSTFMGQVPELARSVCICTELCISYDCV 434

Db 288 LVCKRTE--NPFLRALRLRNQTLTLTSLNMLHSRGLAEVFGILARHN--ISVDLI 342

QY 435 ATSEVSVSLDPSKISR--ELIQASELDHVEEELKAIYVRLIQSAIISLIGN-VE 491

Db 343 TTSEVSVALLTMDTGSTSTGDTLLTQS-----LIMELSLACRVEVEEGALVALIGNDLS 397

QY 432 QSSLILEKTRGVLRKSGVNVQMISQASKYNMGLIYHSDPAKLYEALHQAFFE 545

Db 398 KACGVKEVFGVLEP--FNIRMLCYGASHNLGFLVPGDAEKVQKLNHNLF 449

RESULT 3

US-08-950-737-2

/ Sequence 2, Application US/08950737

/ Patent No. 5932453

/ GENERAL INFORMATION:

/ APPLICANT: KIKUCHI, YOSHIMI

/ APPLICANT: NAKAMISHI, KAZUO

/ APPLICANT: KOJIMA, HIROYUKI

/ TITLE OF INVENTION: PROCESS FOR PRODUCING L-AMINO ACID

/ NUMBER OF SEQUENCES: 8

/ CORRESPONDENCE ADDRESSES:

/ ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,

/ STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400

/ CITY: ARLINGTON

/ STATE: VA

/ COUNTRY: USA

/ ZIP: 22202

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patent In Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/950,737

/ FILING DATE: 15-OCT-1997

/ CLASSIFICATION:

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: JP 272114/1996

/ FILING DATE: 15-OCT-1996

/ ATTORNEY/AGENT INFORMATION:

/ NAME: OBLON, NORMAN F.

/ REGISTRATION NUMBER: 24,618

/ REFERENCE/DOCKET NUMBER: 10-888-0

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (703) 413-3000

/ TELEFAX: (703) 413-2220

/ INFORMATION FOR SEQ. ID NO: 2:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 449 amino acids

/ TYPE: amino acid

/ TOPOLOGY: linear

/ MOLECULE TYPE: protein

US-08-950-737-2

Query Match 20.2%; Score 569; DB 2; Length 449;

Best Local Similarity 34.0%; Pred. No. 1.2e-48;

Matches 161; Conservative 89; Mismatches 182; Indels 42; Gaps 14;

QY 85 VVWKFGSSSVSAARAAVAGLITPPEERPVVLSAMGKTNNLLAGEKAVCGVIHV 144

Db 5 VVSKFGSTVADPDANRSADIVLSDANVR-LVLSASAGITNLVALAE-----GLEPG 58

QY 145 SEIEEMNWKSLHIKIVDELGLPKICNTSLYEQLKGIAMKE-----LTPRTSDYLV 199

Db 59 EREKLDARNIQFALLERLRYP--NVIREIERLENTIVLAALALATSPALTDLV 115

QY 200 SFGECMSTRIFSAIYANKIRVAKROYDAFDIGFITTDEFGAND-----ILATYPAVAKRL 254
 DB 116 SHEELMSTLLFEIILREDFVQAWFDRKV-WRTNDRFGRAEPDIAALAEALALQILPRL 174
 QY 255 HGDWIDOPAIPIVVTGLGKMGSGAVTTLGRGSSDLTATTIGKALGRIQWMDYDGV 314
 DB 175 -----NEGVLITQGFISENK-GRITTLGRGSDYTALALAEALHARSVDIWDVGIY 227
 QY 315 TCDENIYPAKTVPYLTTEBATELAYFGAQLVHPSMRPAREGDIPIRVKNSYNPKAPGT 374
 DB 228 TTDPRVSAKRIIDEIAFAEAEMATFGAKVLPATLLPAYMSDIPVFGSSKDPRAAGT 287
 QY 375 LITRQDMGXGLVLTSLVKSNTVMDIVSTRMIGQYGFARVSGICYIEDLCISYDCV 434
 DB 288 LVGNKTE--NPELFRALALRNQTLTLHSLNMLHSGFLAEVFGILARNH--ISVDLI 342
 QY 435 ATSEVSVSLDPSKIMSR--ELIQASELDHVEELEKIAIVRLLOQRAIISLIGN-VE 491
 DB 343 TISEVVALTLDTGSTGTGDTLLTQS-----LIMEALSALCEVEVEGALVALIGNDLS 397
 QY 492 QSLILEKTRGVLRKSGVNVQMSQASAKVMSLIVHSDAKALVEALHQAFFE 545
 DB 398 KACGVKEVGVLEP--FNIRMTCYGASSHNLCLVPEDEAQVQKLSNLF 449

RESULT 4
 US-08-973-461A-8
 ; Sequence 8, Application US/08973461A

; GENERAL INFORMATION:
 ; APPLICANT: KOJIMA, HIROYUKI
 ; APPLICANT: OGAWA, YURI
 ; APPLICANT: KANAMURA, KAZUE
 ; APPLICANT: SANO, KONOSUKE
 ; TITLE OF INVENTION: METHOD OF L-LYSINE BY FERMENTATION
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT, P.C.
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/973,461A
 ; FILING DATE: 20-APR-1998
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 7-146054
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: NORMAN F. OBLON
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 0010-0901-0 PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-413-3000
 ; TELEFAX: 703-413-2220
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 449
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-973-461A-8

Query Match 20.2%; Score 569; DB 2; Length 449;
 Best Local Similarity 34.0%; Pred. No. 1,2e-48;

Matches 161; Conservative 89; Mismatches 182; Indels 42; Gaps 14;
 QY 85 VMKRGSSSVSSAARMAEVACILITFPEERVVVLSAMGKTTNNLLAGEKAVGCCYIVH 144
 DB 5 VVSKFGGTSVADFDAMNSADIVLSDANVR-LVLLSASAGITNLLVLAEE-----GLEFG 58
 QY 145 SEIEEMNWKSLHIXTVDELQPYXCNTSLYEQLKGIAMKE-----LTPRTSDYLV 159
 DB 59 ERFELDAIRNIGFALIRLRYP---NVIREFIRLLENITVLAEEAALATSPALTDENV 115
 QY 200 SFGECMSTRIFSAIYANKIRVAKROYDAFDIGFITTDEFGAND-----ILATYPAVAKRL 254
 DB 116 SHEELMSTLLFEIILREDFVQAWFDRKV-WRTNDRFGRAEPDIAALAEALALQILPRL 174
 QY 255 HGDWIDOPAIPIVVTGLGKMGSGAVTTLGRGSSDLTATTIGKALGRIQWMDYDGV 314
 DB 175 -----NEGVLITQGFISENK-GRITTLGRGSDYTALALAEALHARSVDIWDVGIY 227
 QY 315 TCDENIYPAKTVPYLTTEBATELAYFGAQLVHPSMRPAREGDIPIRVKNSYNPKAPGT 374
 DB 228 TTDPRVSAKRIIDEIAFAEAEMATFGAKVLPATLLPAYMSDIPVFGSSKDPRAAGT 287
 QY 375 LITRQDMGXGLVLTSLVKSNTVMDIVSTRMIGQYGFARVSGICYIEDLCISYDCV 434
 DB 288 LVGNKTE--NPELFRALALRNQTLTLHSLNMLHSGFLAEVFGILARNH--ISVDLI 342
 QY 435 ATSEVSVSLDPSKIMSR--ELIQASELDHVEELEKIAIVRLLOQRAIISLIGN-VE 491
 DB 343 TISEVVALTLDTGSTGTGDTLLTQS-----LIMEALSALCEVEVEGALVALIGNDLS 397
 QY 492 QSLILEKTRGVLRKSGVNVQMSQASAKVMSLIVHSDAKALVEALHQAFFE 545
 DB 398 KACGVKEVGVLEP--FNIRMTCYGASSHNLCLVPEDEAQVQKLSNLF 449

RESULT 5

US-08-648-010-8

; Sequence 8, Application US/08648010

; Patent No. 6040160

; GENERAL INFORMATION:

; APPLICANT: KOJIMA, HIROYUKI

; APPLICANT: OGAWA, YURI

; APPLICANT: KANAMURA, KAZUE

; APPLICANT: SANO, KONOSUKE

; TITLE OF INVENTION: METHOD OF PRODUCING L-LYSINE BY

; FERMENTATION

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,

; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

; CITY: ARLINGTON

; STATE: VA

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/648,010

; FILING DATE: 29-MAY-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/JP94/01994

; FILING DATE: 28-NOV-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 5/308397

; FILING DATE: 08-DEC-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: OBLON, NORMAN F.

; REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 10-805-0 PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 449 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-648-010-8

Query Match 20.1%; Score 569; DB 3; Length 449;
 Best Local Similarity 34.3%; Pred. No. 1,2e-48;
 Matches 161; Conservative 89; Mismatches 182; Indels 42; Gaps 14;

QY 85 VMKFGSSVSSAARMAEVAAGILTFPEERPVVLSAMGKTTNNLLAGEKAVGCVIHY 144
 DB 5 VVSXFGTGVADPDMNRSADIVLSDANVR-LVVISASAGITNLLVALAE-----GLEPG 58
 QY 145 SEIEEMNWKSLHIKTVDLGPXICNTSLVEBQLKGIAMKE-----LTPRTSDYLV 139
 DB 59 EPEKLDALRNIGFALIERLRYF--NVIREEIERLBNITVLEBAALATSPALTDLVL 115
 QY 200 SPGEQSTRIFSAVYLNKIRVKARQYDAFDIGFITDEFGND-----ILEATYPAVAKRL 254
 DB 116 SHGELNSTLTFEILREKRGISADWFDVAKV-RTINDRFGRAEPDIAALAEIALQLPRL 174
 QY 255 HGDWIDPPALPVVTGFLGKMGSGAVTTLGRGSDLTATTIGKLGRLGIOWKDVGV 314
 DB 175 -----NEGVLITGGIGSENR-GRITTLGRGSDYVALLAEALHMSRVDIWTVDVGIT 227
 QY 315 TCDPNITPAKTVPIYLFEBATELAYFGAQLVHPQSMRPARBGDIPIRVKNSYNPKAGT 374
 DB 228 TIDPRVSAAKRIDELAFEAABEMATFGAKVLHPATLLPAVRSDIPVVGSKCPRAGGT 287
 QY 375 LITRQDMXGLVLTSLVLSKSNVTMLDVTSTMLGQYGFARVSGICYIEDLCISVDC 434
 DB 288 LVCKNKE--NPELFPAALARNQTLTLHSLMLHSGFAEFGLARHN--ISVDLI 342
 QY 435 ATSEVSVSLDPSKIWSR--ELIQASELDHVELEKIALVRLQORALISLIGN-VE 491
 DB 343 TISEVSVALTIDTGTSTGDTLLQS-----LIMELALCRVEBEGIALVALIGDLS 397
 QY 492 QSSLLIEKTRVLRKSGVAVQWISQASKNMSLIVHSDAKALVEALHQAFF 545
 DB 398 KACGVKEVGVLEP--FNIRMIQYASGSHLCLFVPGDAEVOVCKLHSLNLF 449

RESULT 6
 US-09-543-681A-7248
 Sequence 7248, Application US/09543681A
 Patent No. 6605709
 GENERAL INFORMATION:
 APPLICANT: GARY BRETON
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709, 1002-001
 CURRENT APPLICATION NUMBER: US/09/543,681A
 CURRENT FILING DATE: 2000-04-05
 PRIOR APPLICATION NUMBER: US 60/128,706
 PRIOR FILING DATE: 1999-04-09
 NUMBER OF SEQ ID NOS: 8344
 SEQ ID NO 7248
 LENGTH: 460
 TYPE: PRT
 ORGANISM: Proteus mirabilis
 US-09-543-681A-7248

Query Match 20.1%; Score 565; DB 4; Length 460;
 Best Local Similarity 34.3%; Pred. No. 3.2e-48;
 Matches 163; Conservative 82; Mismatches 182; Indels 48; Gaps 15;

QY 86 VMKFGSSVSSAARMAEVAAGILTFPEERPVVLSAMGKTTNNLLAGEKAVGCVIHY 145
 DB 18 IAKFGGTSVANFAMEKCAIILKQSVR-VVLSASAGITNLLI-----ELATVEARQ 71
 QY 146 EIEEMNWKSLHIKTVDLGPXICNTSLVEBQLKGIAMKE-----LTPRTSDYLV 200
 DB 72 RAALLAQVRIEVALINQISQAEIIS--QEINRLLENITEMSEAAALATSPALTDLVL 128
 QY 201 FGEQSTRIFSAVYLNKIRVKARQYDAFDIGFITDEFGND-----LLEATYPAVAKR 253
 DB 129 HGEIINSTLTF--VELIREKIGISADWFDVAKVKTNDLPCHAPDMAQLTELTQSIQCR 185
 QY 254 LHGDWIDPPALPVVTGFLGKMGSGAVTTLGRGSDLTATTIGKLGRLGIOWKDVGV 313
 DB 186 L-----TEVITVGGTIGGEPK-GKITTLGRGSDYVALLAEALGMSRVIDIWTVDGSI 238
 QY 314 LTPDPNITPAKTVPIYLFEBATELAYFGAQLVHPQSMRPARBGDIPIRVKNSYNPKAG 373
 DB 239 YSTDPRIYQAHRIIDHIAFDEAEMATEFGAKILHPATLLPAVRSGIPVVGSKAPBEG 298
 QY 434 VATSEVSVSLDPSKIWSR--ELIQASELDHVELEKIALVRLQORALISLIGN-- 489
 DB 354 ITTSVSVSIALTIDTGTSTSSGSLITNA-----LTELALCRVEEEDLAVAILGNL 408
 QY 490 VEGSLLIEKTRVLRKSGVAVQWISQASKNMSLIVHSDAKALVEALHQAFF 544
 DB 409 SQVNLGSGITGTL---EDNIRMIISHGASTHNLCLLVGCKADNIVRKHLDTLF 460

RESULT 7
 US-09-489-039A-8740
 Sequence 8740, Application US/09489039A
 Patent No. 6610836
 GENERAL INFORMATION:
 APPLICANT: Gary Breton et. al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709, 2004001
 CURRENT APPLICATION NUMBER: US/09/489,039A
 CURRENT FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: US 60/117,747
 PRIOR FILING DATE: 1999-01-29
 NUMBER OF SEQ ID NOS: 14342
 SEQ ID NO 8740
 LENGTH: 835
 TYPE: PRT
 ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-8740

Query Match 16.7%; Score 469.5; DB 4; Length 835;
 Best Local Similarity 29.1%; Pred. No. 3.8e-38;
 Matches 141; Conservative 101; Mismatches 207; Indels 35; Gaps 13;

QY 86 VMKFGSSVSSAARMAEVAAGILTFPEERPVVLSAMGKTTNNLLAGEKAVGCVIHY 144
 DB 18 VIKFGGTSVANFAMEKCAIILKQSVR-VVLSASAGITNLLI-----ELATVEARQ 71
 QY 145 SEIEEMNWKSLHIKTVDLGPXICNTSLVEBQLKGIAMKE-----LTPRTSDY 137
 DB 77 PNIAAERIFAEILLOGLADAPAPFLAQLKAFVQBEFAQIKVHLHGLSLGQCPDSVNA 136
 QY 198 LVSPGQSTRIFSAYL---NKIRVKARQYDAFDIGFITDEFGNDILEATYPAVAKR 253
 DB 137 LICRGKSLIAMAELARHGKSVINPVKLAAGHYLE---STVDIAESTRRRLAASQ 193
 QY 254 LHGDWIDPPALPVVTGFLGKMGSGAVTTLGRGSDLTATTIGKLGRLGIOWKDVGV 313
 DB 194 IPAD-----HMTIMAGFTA-GNEKGELVVLGRNSDYSAAVLAACLRADCCETIWTVDGV 247

Db 219 ALNIPKDHMLINAGFTA-GNEKNELVILGRNGSDYSAVAALACLAQCCEIWTVDVGYT 277
 QY 316 CDPNITPAKTVPYLTFFEEATLTAAGVYLHPQSMRPARBEDIPIVRKNSYNKAGTL 375
 Db 278 CDRILVPAHLLKGRSFGAMELSYFGAKVLPRTIAPACQIPCLIKNGPEAEGTL 337
 QY 376 I-TRORDMXGLVLTSLVILKSNVTMLDIVSTRMLGOYGFARVSGICYIEDLCISYDCV 434
 Db 338 IGGGQKDDSTPVYGIINL---NNMAMINVSQPMKGMVGMARVPSVMSRAGISVLTQL 394
 QY 435 ATSEVSVSLDPSKTSRELLIOQASEL-----DHVELEKIAIVLLQORAIIS 485
 Db 395 SSSEVSIISFCV-PQK---ELIRPAQALSEEFYELKOGVLDPLD-----IMNNVALIS 443
 QY 486 LIQN-VEQSSLIETKGTVRKSGVNVQMSQASKVNMSLIVHSDAKALVEALHQAFF 544
 Db 444 VVDDGMRILKGIARFPAITRGNITVAIAGSSERSISAVIANDATNVRLOHQLF 503
 QY 545 EDDVLSQV 552
 Db 504 NTQIVEV 511
 RESULT 10
 US-09-107-532A-6090
 / Sequence 6090, Application US/09107532A
 / Patent No. 6583275
 / GENERAL INFORMATION:
 / APPLICANT: Lynn A Doucette-Stamm and David Bush
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 / ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 / NUMBER OF SEQUENCES: 7310
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 / STREET: 100 Beaver Street
 / City: Waltham
 / STATE: Massachusetts
 / COUNTRY: USA
 / ZIP: 02354
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: CD-ROM ISO9660
 / COMPUTER: PC
 / OPERATING SYSTEM: <Unknown>
 / SOFTWARE: ASCII
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/107,532A
 / FILING DATE: 30-Jun-1998
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: 60/085,598
 / FILING DATE: 14 May 1998
 / APPLICATION NUMBER: 60/051571
 / FILING DATE: July 2, 1997
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Arianello, Pamela Deneka
 / REGISTRATION NUMBER: 40,469
 / REFERENCE/DOCKET NUMBER: GTC-012
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (781)893-8277
 / TELEFAX: (781)893-8277
 / INFORMATION FOR SEQ ID NO: 6090:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 478 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / MOLECULE TYPE: protein
 / HYPOTHEICAL: YES
 / ORIGINAL SOURCE:
 / ORGANISM: Enterococcus faecium
 / FEATURE:
 / NAME/KEY: misc_feature
 / LOCATION: (B) LOCATION 1...478
 / SEQUENCE DESCRIPTION: SEQ ID NO: 6090:
 US-09-107-532A-6090

Query Match 15.2%; Score 428.5; DB 4; Length 478;
 Best Local Similarity 27.2%; Pred. No. 2.1e-34;
 Matches 130; Conservative 96; Mismatches 193; Indels 59; Gaps 15;
 QY 86 VMKFGSSVSGAARMAEYAGLITFFEEPRPVVLSAMGKTNNLLAGEAVACGYI--- 142
 Db 32 VAKFGTSMHSEDEPKYRTIVLSDP-NROVIVSALGKRTSD---DDKVTDLVLYIA 86
 QY 143 ----HVEIEEMNMY--KSLIKTVDELGLPXICNTSLVEYELQLKGIAMMKELTPRTSD 196
 Db 87 YLGHVDMVTPMKRCDFRIGRVVLAQLP-----IDELARIQKQLEDEKITED 137
 QY 197 YVVSFGECMSTRISAYINKIRVARQYDAFDIGFITTFEGNADILEA--TPYAVAKL 254
 Db 138 YVVSFGVYLTQALMAEYLYQFIDAKLIFEDYDGKIDOKRTAOELTKARXPYK--- 193
 QY 255 HGDWQDPAIPVYVYGLGKWSGAVTTTLGSGSDLTATITGKALGLREIQVWKDVGVL 314
 Db 194 -----VPGFYGAN-PAGEKELGSGSDITGALLATLQAESYEMTVDYSGIM 241
 QY 315 TCDPNIYPKATVPYLTFFEEATLTAAGVYLHPQSMRPARBEDIPIVRKNSYNPAPGT 374
 Db 242 MADPRLIDHPKRIEISFRELEMAWGANVLEAAVPEVQEDIP:QIKNTDPPAOGT 301
 QY 375 LTRQR-DMDXGLVLTSLVILKSNVTMLDIVSTRMLGOYGFARVSGICYIEDLCISYDC 433
 Db 302 KISNHEIEKENG--LTSLAKRDPFLITLTKRAMSDEIGFIKMSIFPKHE--ISIEH 356
 QY 434 VATSEVSVSLDPSKTSRELLIOQASELDHVELEK---IAIVLLQORAIISLIGNV 490
 Db 357 IESGIDNIGV-----IVSAEAL--ADRLFITREIKETLGEIEIVIEDLALISVYGGP 408
 QY 491 EGSLLIE-KTGRVYARKGVNVQMSQASKVNMSLIVHSDAKALVEALHQAFFED 547
 Db 409 HKEILGSGKVLSTLNKLEIRTSLSQCAQELNLTIGVNNQIVETVVKGIYEGWNTD 466
 RESULT 11
 US-08-380-182-19
 / Sequence 19, Application US/08380182
 / Patent No. 5858749
 / GENERAL INFORMATION:
 / APPLICANT: Matthews, Benjamin F.
 / TITLE OF INVENTION: A Bifunctional Protein From Carrots
 / TITLE OF INVENTION: (Daucus carota) with Aspartokinase and Homoserine
 / TITLE OF INVENTION: Dehydratogenase Activities
 / NUMBER OF SEQUENCES: 23
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Janelle S. Graeter
 / STREET: Room 411, Bldg. 005, BARC-W
 / City: Beltsville
 / STATE: Maryland
 / COUNTRY: USA
 / ZIP: 20705
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: PatentIn Release #1.0, Version #1.25
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/380,182
 / FILING DATE:
 / CLASSIFICATION: 435
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Graeter, Janelle S.
 / REGISTRATION NUMBER: 35,024
 / REFERENCE/DOCKET NUMBER: 0226.94
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 301-504-6629
 / TELEFAX: 301-504-5060
 / INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:
 LENGTH: 863 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-380-182-19

Query Match 14.8%; Score 416.5; DB 2; Length 863;
 Best Local Similarity 26.3%; Pred. No. 9, 2e-33;
 Matches 126; Conservative 108; Mismatches 210; Indels 35; Gaps 16;

QY 86 VMKFGSSVSSAARMAEVAAGLITFPEERPVVLSAMGKTIN--NLLAGEKAVGCGVI 142
 DB 36 IHKFGTGVSSSRIRHVAEIVVEDSEKRLVVSAMSQVTDMDLTKAKSRDSEY 95
 QY 143 HVEIEEMNVKSLHKTVDLG-LEPICNTSLVELEQLKGIAMKELTPTSDPLVSF 201
 DB 96 ALDAAVEKHKLTFDLDDGLARFLRLQHDVNNKALRAIYAGHATESFSDPVGH 155
 QY 202 GECSSTRIFSAVYLNKIRYARQYDAFDIGITTFDEFGNADLEACTYPAVAKRLHGDWIQD 261
 DB 156 GEIWSAQLLSFVIRKGGDCNMMDTRDY--LVNPAQSNQV-DPDYLESEKRLK-KWSS 211
 QY 262 PAIP--VTGFLGKSKSAVITLGGSDLTATTIIGKALREIQWKVDVDTGVLTCOPN 319
 DB 212 NQCQITVATGFI-FTPCNIPTTLKRDGSDFSALIMGALLRAGQVITWTVNGVYSADPR 270
 QY 320 IYPAKTVPLTFEEATELAYFGAQLVHPQSMRPAEGDIPVRKNSYNPAKPGTLLTRQ 379
 DB 271 KVEEAVVYKTLSSQEMWESYFGANVLRHTIIPVVRVYDIPVININFLSAPGTMCRE 330
 QY 380 R--DMDXGLV--LTSIVLKSNTMLDIVSTRMLGOYFLARVSGICIEDLCISVDCV 434
 DB 331 SVGETEDGLTESHVGFATIDMLALINEGTGMAGVPGTAIGA--VKDVGANVMI 388
 QY 435 --ATSEVSVSLDPSKISRELIIQASE-----LDHVELEKIAIVLLQORATISL 486
 DB 389 SQASSEHSICFAVPESEV--KAAVALARPQALD--ANLSQVALI--PNCSTIAT 440
 QY 487 IGNEOSGLILEKT-GRVLRKSGVNVQMTSQGASKYNSLIYHSDAKALVEALHQAFF 544
 DB 441 VQKMASTPGVSAITLFWALAKANINRAIAQCTEYINIVLSREDCVARALAVHSREY 499

RESULT 12
 US-09-134-001C-4160
 / Sequence 4160, Application US/09134001C
 / Patent No. 6380370
 / GENERAL INFORMATION:
 / APPLICANT: Lynn Doucette-Stamm et al
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 / TITLE OF INVENTION: EPIDERMIDS FOR DIAGNOSTICS AND THERAPEUTICS
 / FILE REFERENCE: GTC-007
 / CURRENT APPLICATION NUMBER: US/09/134,001C
 / CURRENT FILING DATE: 1998-08-13
 / PRIOR APPLICATION NUMBER: US 60/064,964
 / PRIOR FILING DATE: 1997-11-08
 / PRIOR APPLICATION NUMBER: US 60/055,779
 / PRIOR FILING DATE: 1997-08-14
 / NUMBER OF SEQ ID NOS: 5674
 / SEQ ID NO 4160
 / LENGTH: 463
 / TYPE: PRT
 / ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-4160

Query Match 14.8%; Score 415.5; DB 4; Length 463;
 Best Local Similarity 27.0%; Pred. No. 4, 2e-33;
 Matches 127; Conservative 108; Mismatches 205; Indels 31; Gaps 13;

QY 80 GDQLSVMKFGSSVSSAARMAEVAAGLITFPEERPVVLSAMG-----KTNMLLA 132
 DB 2 GRKIMYAKFGSSVSTABQIKVLTIVNEDP-ERKIIIVSAPGRKHNDDIKTTDLRL 60

QY 133 GEKAVGCGVIHVEIEEMNVKSLHKTVDLGLEPICNTSLVELEQLKGIAMKELT 191
 DB 61 YEKVLN-KLWYSSKQCE---IIQRYADIVELG---IGNDLININPTLEBYIKHLSKP 113
 QY 192 PRTSDVLSFGECSSTRIFSAVYLNKIRYARQYDAFDIGITTFDEFGNADLEACTYPAVA 251
 DB 114 NRYDALISCGENFNAQIAQYNNSSOGIPTEYISPKRAGLVTVLPOOAQILLSAVNEIY 173
 QY 252 KALHGDWIDPAIPVTVFLGKWSGAVITLGGSDLTATTIIGKALREIQWKVDV 311
 DB 174 KL--RDYDEKLIIP--GPGVVS-KQNYIVTFPPGSGDITATIIARGVRSVLYNFTDVS 227
 QY 312 GVLTCOPNIYPAKTVPLTFEEATELAYFGAQLVHPQSMRPAEGDIPVRKNSYNPAK 371
 DB 228 GIYKANPNIINNPELIEETVREMSELISVAGFGVHDEALQPLKCRIPVVIKNTNRPND 287
 QY 372 PGTLITRQDMXGLVLTSLVLSNTVMDIVSTRMLGOYFLARVSGICIEDLCISV 431
 DB 288 KGTYLHDEEIDSKNVI-SGISCDKGTVINIKYIKNRLVGFTRKILGV--LEEFNISF 344
 QY 432 DCVATSEVSVSLDPSKISRELIIQASELDHVELEKIAIVLLQORATISLIG-NV 490
 DB 345 DHMPSGIDNISITMETNOIQKE-----SQVLNARIRCEVDELSIDHDLAVIMTIVEGM 359
 QY 491 EQSGLILEKTVLRKSGVNVQMTSQGASKYNSLIYHSDAKALVEALHQA 541
 DB 400 NVVGTASKITHALSESINILIMINQASSELISWFGIHEADAKAVLSTYE 450

RESULT 13
 US-08-532-828B-3
 / Sequence 3, Application US/08532828B
 / Patent No. 5686671
 / GENERAL INFORMATION:
 / APPLICANT: SUGIMOTO, Masakazu
 / APPLICANT: OGAWA, Yuki
 / APPLICANT: SUZUKI, Tomoko
 / APPLICANT: TANAKA, Akiko
 / APPLICANT: WATSUI, Hiroshi
 / TITLE OF INVENTION: MUTANT ASPARTOKINASE GENE
 / NUMBER OF SEQUENCES: 24
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
 / STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 / CITY: ARLINGTON
 / STATE: VA
 / COUNTRY: USA
 / ZIP: 22202
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: MS-DOS TEXT EDITOR
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/532,828B
 / FILING DATE: 27-OCT-1995
 / CLASSIFICATION: 435
 / PRIOR APPLICATION DATA:
 / PRIOR APPLICATION NUMBER: JP 5-101450
 / FILING DATE: 27-APR-1993
 / ATTORNEY/AGENT INFORMATION:
 / NAME: NORMAN F. OBLON
 / REGISTRATION NUMBER: 24,618
 / REFERENCE/DOCKET NUMBER: 10-764-0 PCT
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 703-413-3000
 / TELEFAX: 703-413-2220
 / INFORMATION FOR SEQ ID NO: 3:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 421 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear

MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC13639
US-08-532-828B-3

Query Match
Best Local Similarity 14.7%; Score 414.5; DB 1; Length 421;
Best Local Similarity 28.5%; Pred. No. 4.5e-33;
Matches 135; Conservative 83; Mismatches 169; Indels 87; Gaps 14;

85 VYMKFGSSVSSAARMAEVALGLITLPEE--ERPVVVLSAMGKTTNNLLAGEKAVGCVI 142
4 VQKYGSSLSAEIRVNAERIVATKAGNDVYVCSAGDITDEL-----51
143 HVSEIEMNMVNSLHIKTVDLGLPXICNTSLYELEQLIKGIAMMKELTPRTSDYLSFG 202
52 -----ELAAVNPVPPAREM-----DMLLTAG 73
203 ECMSTRIFSAVINKIRVAKROYDAPDIGITTFDEFGNADILEATYPVAKRLHGMIDP 262
74 ERISNALVMAIESIGAEKQSFQAGVLTTERHGAKRIVDTTPGRVREAL-----DEG 128
263 AIPVVTGFLGKWKSGAVTTLGRGSDLTATTIGKALGREIQWKVDVGLTCDPNIYP 322
129 KICIVAGFQGVNKEKTRDVTTLGRGSDTTAVALAALNADVCEIYSDVDGYTADPRIVP 188
323 HAKTVPYLTPEEATLAFGAOVLHPQSMRPARBGDIPIRVKNSINPAPRTIIT-RQRD 361
189 NAQKLEKLSFEEMLELAAGSKILVRSVEYARAFNPLRVRSYS-NDPGLIAGSMED 247
382 MDXGLVLTSTIVL--KSNVTMLDIYSTRLMGQYGLFARVSGICYIEDLCISVDCVATSE 438
248 IVEEAVLTGVAITDKSEAKVTVLGISD-----KPGETAKV--FRALDAEINIDMTVLQNV 300
439 VSVSVLSDPSKIMSELIQQAISELD--HVEELEKIAL-----YRLQQRALISLIG-N 489
301 SSVEDG-----TTDITFTCPRADGRRAHEILKLVQGVNWTNVLVDYDQVKXSLVAG 353
490 VEOSSLILEKTRVLRKSGVNVOMISQASKYNNLSLIVHSDAKALVEALHQAIF 543
354 MKSHPGVTAEPFMAELRDVNVNIELIS--TSEIRISVILREDDLDAAARALHGEF 405

RESULT 14
US-08-532-828B-4
Sequence 4, Application US/08532828B
Patent No. 5688671
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: OGAWA, Yumi
APPLICANT: SUZUKI, Tomoko
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
TITLE OF INVENTION: MUTANT ASPARTOKINASE GENE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS TEXT EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532, 828B
FILING DATE: 27-OCT-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-101450
FILING DATE: 27-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-764-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 421 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: AJ3463
US-08-532-828B-4

Query Match
Best Local Similarity 14.7%; Score 414.5; DB 1; Length 421;
Best Local Similarity 28.5%; Pred. No. 4.5e-33;
Matches 135; Conservative 83; Mismatches 169; Indels 87; Gaps 14;

85 VYMKFGSSVSSAARMAEVALGLITLPEE--ERPVVVLSAMGKTTNNLLAGEKAVGCVI 142
4 VQKYGSSLSAEIRVNAERIVATKAGNDVYVCSAGDITDEL-----51
143 HVSEIEMNMVNSLHIKTVDLGLPXICNTSLYELEQLIKGIAMMKELTPRTSDYLSFG 202
52 -----ELAAVNPVPPAREM-----DMLLTAG 73
203 ECMSTRIFSAVINKIRVAKROYDAPDIGITTFDEFGNADILEATYPVAKRLHGMIDP 262
74 ERISNALVMAIESIGAEKQSFQAGVLTTERHGAKRIVDTTPGRVREAL-----DEG 128
263 AIPVVTGFLGKWKSGAVTTLGRGSDLTATTIGKALGREIQWKVDVGLTCDPNIYP 322
129 KICIVAGFQGVNKEKTRDVTTLGRGSDTTAVALAALNADVCEIYSDVDGYTADPRIVP 188
323 HAKTVPYLTPEEATLAFGAOVLHPQSMRPARBGDIPIRVKNSINPAPRTIIT-RQRD 361
189 NAQKLEKLSFEEMLELAAGSKILVRSVEYARAFNPLRVRSYS-NDPGLIAGSMED 247
382 MDXGLVLTSTIVL--KSNVTMLDIYSTRLMGQYGLFARVSGICYIEDLCISVDCVATSE 438
248 IVEEAVLTGVAITDKSEAKVTVLGISD-----KPGETAKV--FRALDAEINIDMTVLQNV 300
439 VSVSVLSDPSKIMSELIQQAISELD--HVEELEKIAL-----YRLQQRALISLIG-N 489
301 SSVEDG-----TTDITFTCPRADGRRAHEILKLVQGVNWTNVLVDYDQVKXSLVAG 353
490 VEOSSLILEKTRVLRKSGVNVOMISQASKYNNLSLIVHSDAKALVEALHQAIF 543
354 MKSHPGVTAEPFMAELRDVNVNIELIS--TSEIRISVILREDDLDAAARALHGEF 405

RESULT 15
US-08-700-359-9
Sequence 9, Application US/08700359
Patent No. 5766925
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, MASAKAZU
APPLICANT: USUDA, YOSHITIRO
APPLICANT: SUZUKI, TOMOKO
APPLICANT: TANAKA, AKIKO
APPLICANT: MATSUI, HIROSHI
TITLE OF INVENTION: METHOD OF PRODUCING L-LYSINE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS TEXT EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,359
FILING DATE: 08-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-35019
FILING DATE: 04-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-819-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-2220
TELEFAX: 703-413-3000
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 421
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-700-359-9

Query Match 14.7%; Score 414.5; DB 1; Length 421;
Best Local Similarity 28.5%; Pred. No. 4.5e-33;
Matches 135; Conservative 83; Mismatches 169; Indels 87; Gaps 14;

QY 85 VVKKFGSSSVSSAARMAEVAAGLITPE--ERYVVISAMGKTTNNLLAGEKAVGCGVI 142
DB 4 VVCKYGGSSLESARIRINVAERIVATKAGNDVYVCSAMGDTTDEL----- 51
QY 143 HVSEIEMNMVKS.HITVDELGLPXICNTSLVELEQLKIMMKELTPRTSDYLVFG 202
DB 52 -----ELAAVNPVPAKEM-----DMLLTNG 73
QY 203 EGMSTRIFSAVINKIRYKARQYDAFDIGITTDGFGNADILEATYPAVAKRLHGWIOGP 262
DB 74 ERISNALVAMAEISLGAESFTGSCAGVLTTERHGNARIYDVTGPRVREAL----DEG 128
QY 263 AIPVVTGFLGKMGSGAVTLTGRGSDLTATTIGKALGLEIQVMKDVGLTCDPNTYE 322
DB 129 KICIVAGQGVNKAETRVTLTGRGSDTLTVALAALNADVCEIYSDVDGVTADPRIVE 188
QY 323 HAKTVPYLTFEATELAVFGAQLHPQMRPAREGDIPIYVKNKSNPKAPGTLIT--RORD 381
DB 189 NAGLEKLSFEEMELIAVSGKILVRSVEYARAPVPLKVRSSYS--NDPGTLIAGSMED 247
QY 382 MDKGLVLTSLVL--KSNVTMLDIVSTRMLGOYGFAPARVSGICIEDLCISVDCVATSE 438
DB 248 IPVEEAVLTGATDKSRAKTVLGISD---KPEGAKV--PRALDAEINIDVYLQNV 300
QY 439 VSVSVSLDPSKIMRELIQASELD--HVEEELKIAI-----VRLLOQRAIISLIG-N 489
DB 301 SSVEDG-----TTDTITFCPRADGRAMEILKLGQGNWNTVLYDDQGVKSLVGAG 353
QY 490 VEOGSLIEKTVLRKSGVNVQVMTISQASKANSLIVHSDAKALVZALHOAF 543
DB 354 MKSHPGVTAEFMEALRDVNVNIELIS--TSEIRISVLIREDDDAAARALHROF 405

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:17:42 ; Search time 67 Seconds
(without alignments)

2164.403 Million cell updates/sec

Title: US-09-890-813-6

Perfect score: 2813

Sequence: 1 MAIPVSAAPRLVPSIRP.....QAFEDVLQVEANILLVG 560

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 25895539 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	2809	99.9	560	US-09-890-813-6	Sequence 6, Appli
2	2678.5	95.2	555	US-09-890-813-8	Sequence 8, Appli
3	1981.5	70.4	415	US-10-425-114-50163	Sequence 50163, A
4	1839	65.4	564	US-10-424-599-170907	Sequence 170907, A
5	1835	65.2	564	US-09-890-813-17	Sequence 17, Appli
6	1826.5	64.9	493	US-10-425-114-52733	Sequence 52733, A
7	1754.5	62.4	561	US-10-424-599-170902	Sequence 170902, A
8	1629.5	57.9	439	US-09-890-813-16	Sequence 16, Appli
9	1587.5	56.4	426	US-10-425-114-70672	Sequence 70672, A
10	1573	55.9	426	US-10-425-114-39779	Sequence 39779, A
11	1222.5	43.5	326	US-10-425-114-39716	Sequence 39716, A
12	992.5	35.3	249	US-10-425-114-43083	Sequence 43083, A
13	972	34.6	221	US-10-425-114-54453	Sequence 54453, A
14	890.5	31.7	281	US-09-890-813-12	Sequence 12, Appli
15	762.5	27.1	226	US-10-425-114-46004	Sequence 46004, A

16	671.5	23.9	464	US-10-369-493-11215	Sequence 11215, A
17	666	23.7	152	US-09-890-813-4	Sequence 4, Appli
18	662	23.5	473	US-10-369-493-21474	Sequence 21474, A
19	642.5	22.8	462	US-10-369-493-978	Sequence 978, App
20	607	21.6	473	US-10-369-493-584	Sequence 584, App
21	604.5	21.5	465	US-10-369-493-8972	Sequence 8972, App
22	586.5	20.8	178	US-10-425-114-51028	Sequence 51028, A
23	578.5	20.6	817	US-10-369-493-10372	Sequence 10372, A
24	575	20.4	447	US-10-369-493-21212	Sequence 21212, A
25	572	20.3	449	US-09-899-339-23	Sequence 23, Appli
26	569	20.2	449	US-10-369-493-23654	Sequence 23654, A
27	552	19.6	448	US-10-369-493-503	Sequence 503, App
28	535	19.0	481	US-10-369-493-10467	Sequence 10467, A
29	508	18.1	811	US-10-282-122A-48795	Sequence 48795, A
30	495.5	17.6	489	US-10-369-493-22563	Sequence 22563, A
31	491.5	17.5	527	US-10-369-493-1741	Sequence 1741, App
32	490.5	17.4	815	US-10-282-122A-66881	Sequence 66881, A
33	483.5	17.2	825	US-10-282-122A-77439	Sequence 77439, A
34	478.5	17.0	580	US-10-282-122A-56011	Sequence 56011, A
35	477.5	17.0	820	US-10-282-122A-42805	Sequence 42805, A
36	475.5	16.9	820	US-10-369-493-686	Sequence 686, App
37	475.5	16.9	815	US-10-282-122A-58074	Sequence 58074, A
38	470.5	16.7	822	US-10-369-493-15294	Sequence 15294, A
39	469.5	16.7	820	US-10-282-122A-59647	Sequence 59647, A
40	469.5	16.7	820	US-10-282-122A-74838	Sequence 74838, A
41	469.5	16.7	820	US-10-282-122A-75391	Sequence 75391, A
42	468.5	16.7	454	US-10-282-122A-73500	Sequence 73500, A
43	466	16.6	818	US-10-369-493-19453	Sequence 19453, A
44	465.5	16.5	812	US-10-369-493-15660	Sequence 15660, A
45	465.5	16.5	812	US-10-369-493-16051	Sequence 16051, A

ALIGNMENTS

RESULT 1
US-09-890-813-6
Sequence 6, Application US/09890813
Publication No. US20020183486A1
GENERAL INFORMATION:
APPLICANT: E. I. du Pont de Nemours and Company
TITLE OF INVENTION: Aspartate Kinase
FILE REFERENCE: BB1430 PCT
CURRENT APPLICATION NUMBER: US/09/890,813
CURRENT FILING DATE: 2001-06-02
PRIOR APPLICATION NUMBER: 60/172294
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Microsoft Office 97
SEQ ID NO 6
LENGTH: 560
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: UNSURE
LOCATION: (168)
NAME/KEY: UNSURE
LOCATION: (384)
US-09-890-813-6
Query Match 99.9%; Score 2809, DB 9, Length 560;
Best Local Similarity 100.0%; Pred. No. 1.4e-267;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAIPVSAAPRLVPSIRPSSGVRGLACGTGTGRGARGLSMVVADTSRRKAD 60
1 MAIPVSAAPRLVPSIRPSSGVRGLACGTGTGRGARGLSMVVADTSRRKAD 60
Db 1 MAIPVSAAPRLVPSIRPSSGVRGLACGTGTGRGARGLSMVVADTSRRKAD 60
QY 61 GGGGVGAPVGLGEGIGDOLSVYMKFGSSVSAAPABVACILITPEEPYVVL 120
61 GGGGVGAPVGLGEGIGDOLSVYMKFGSSVSAAPABVACILITPEEPYVVL 120
Db 61 GGGGVGAPVGLGEGIGDOLSVYMKFGSSVSAAPABVACILITPEEPYVVL 120
121 AMGKTNNLLGKAVGCGVIVHSEIEEMNNVKSILHKTVDGLPXCITSLEYELQ 180

Db 121 AMGKTTNNLLAGEKAVCGVTHVSEIEMNNVKSLLIKTYDELGLPXICNTSLYLEEQL 180
 QY 181 LKGIAMMKELTPRTSDYLVSFGECSMSTRIFSAYNKIRVAKROYDADFQITTFDEFGNA 240
 Db 181 LKGIAMMKELTPRTSDYLVSFGECSMSTRIFSAYNKIRVAKROYDADFQITTFDEFGNA 240
 QY 241 DILEATYPAVARLHGDWIODPAIPVVTGFLGKGMKSGAVTTLLRGSSDLTATTIGKALG 300
 Db 241 DILEATYPAVARLHGDWIODPAIPVVTGFLGKGMKSGAVTTLLRGSSDLTATTIGKALG 300
 QY 301 LREIOWKVDVGLTCDPNIPYHAKTYVYLTFEEATELAYFGAQLHPQSRPARBEDI 360
 Db 301 LREIOWKVDVGLTCDPNIPYHAKTYVYLTFEEATELAYFGAQLHPQSRPARBEDI 360
 QY 361 VARVKSYNKAPGTLITRQDMDXGLVLTSLVLSKNTMLDIYSTMLGQYGLARVSG 420
 Db 361 VARVKSYNKAPGTLITRQDMDXGLVLTSLVLSKNTMLDIYSTMLGQYGLARVSG 420
 QY 421 ICYIEDLCISVDCVATSEVSVSLDPKIMSRELIQOASELDHVEELEKIALVRLQ 480
 Db 421 ICYIEDLCISVDCVATSEVSVSLDPKIMSRELIQOASELDHVEELEKIALVRLQ 480
 QY 481 RAIIISLIGNEOSSLILEKTGRVLRKSGVNVQMISQASKNMSLIYHDSAKALVEALH 540
 Db 481 RAIIISLIGNEOSSLILEKTGRVLRKSGVNVQMISQASKNMSLIYHDSAKALVEALH 540
 QY 541 QAFEDDVLISQVEAENLTVG 560
 Db 541 QAFEDDVLISQVEAENLTVG 560

RESULT 2

US-09-890-813-8
 ; Sequence 8, Application US/09890813
 ; Publication No. US20020183486A1
 ; GENERAL INFORMATION:
 ; APPLICANT: E. I. du Pont de Nemours and Company
 ; TITLE OF INVENTION: Aspartate Kinase
 ; FILE REFERENCE: BB1430 PCT
 ; CURRENT APPLICATION NUMBER: US/09/890, 813
 ; CURRENT FILING DATE: 2001-08-02
 ; PRIOR APPLICATION NUMBER: 60/172944
 ; PRIOR FILING DATE: 1999-12-21
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 8
 ; LENGTH: 565
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; US-09-890-813-8

Query Match 95.2%; Score 2678.5; DB 9; Length 555;
 Best Local Similarity 96.8%; Pred. No. 1e-254;
 Matches 543; Conservative 1; Mismatches 10; Indels 7; Gaps 4;

QY 1 MAIPVRSAAAPRRLLVPSIPPASSGAVRGLACTGRTGPRGAAGLSMVAADSTRRAKAD 60
 Db 1 MAIPVRSAAAPRRLLVPSIPPASSGAVRGLACTGRTGPRGAAGLSMVAADSTRRAKAD 60
 QY 61 GGGGVGAPVGLGLGMEGLDQDLSYVMKFGSSVSAAAMAVAGLILTFPEERVVYLS 120
 Db 61 GGGGVGAPVGLGLGMEGLDQDLSYVMKFGSSVSAAAMAVAGLILTFPEERVVYLS 120
 QY 121 AMGKTTNNLLAGEKAVCGVTHVSEIEMNNVKSLLIKTYDELGLPXICNTSLYLEEQL 180
 Db 121 AMGKTTNNLLAGEKAVCGVTHVSEIEMNNVKSLLIKTYDELGLPXICNTSLYLEEQL 180
 QY 180 LKGIAMMKELTPRTSDYLVSFGECSMSTRIFSAYNKIRVAKROYDADFQITTFDEFGN 239
 Db 181 LKGIAMMKELTPRTSDYLVSFGECSMSTRIFSAYNKIRVAKROYDADFQITTFDEFGN 240
 QY 240 ADILEATYPAVARLHGDWIODPAIPVVTGFLGKGMKSGAVTTLLRGSSDLTATTIGKAL 299

Db 241 ADILEATYPAVARLHGDWIODPAIPVVTGFLGKGMKSGAVTTLLRGSSDLTATTIGKAL 300
 QY 300 GLREIOWKVDVGLTCDPNIPYHAKTYVYLTFEEATELAYFGAQLHPQSRPARBEDI 359
 Db 301 GLREIOWKVDVGLTCDPNIPYHAKTYVYLTFEEATELAYFGAQLHPQSRPARBEDI 360
 QY 360 PVKKSYNKAPGTLITRQDMDXGLVLTSLVLSKNTMLDIYSTMLGQYGLARVSG 419
 Db 361 PVKKSYNKAPGTLITRQDMDXGLVLTSLVLSKNTMLDIYSTMLGQYGLARVSG 418
 QY 420 GICYIEDLCISVDCVATSEVSVSLDPKIMSRELIQOASELDHVEELEKIALVRLQ 479
 Db 419 AI--FEDLCISVDCVATSEVSVSLDPKIMSRELIQ--ELDHVEELEKIALVRLQ 474
 QY 480 QRAIIISLIGNEOSSLILEKTGRVLRKSGVNVQMISQASKNMSLIYHDSAKALVEALH 539
 Db 475 QRAIIISLIGNEOSSLILEKTGRVLRKSGVNVQMISQASKNMSLIYHDSAKALVEALH 534
 QY 540 HQAFEDDVLISQVEAENLTVG 560
 Db 535 HQAFEDDVLISQVEAENLTVG 555

RESULT 3

US-10-425-114-50163
 ; Sequence 50163, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425, 114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 50163
 ; LENGTH: 415
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700553382_F11.pep
 ; US-10-425-114-50163

Query Match 70.4%; Score 1981.5; DB 12; Length 415;
 Best Local Similarity 95.7%; Pred. No. 3.4e-186;
 Matches 403; Conservative 1; Mismatches 10; Indels 7; Gaps 4;

QY 141 VHVSEIEMNNVKSLLIKTYDELGLPXICNTSLYLEEQLKGIAMMKELTPRTSDYLV 139
 Db 1 VHVSEIEMNNVKSLLIKTYDELGLPXICNTSLYLEEQLKGIAMMKELTPRTSDYLV 60
 QY 200 SFGECSMSTRIFSAYNKIRVAKROYDADFQITTFDEFGNADILEATYPAVARLHGDW 259
 Db 61 SFGECSMSTRIFSAYNKIRVAKROYDADFQITTFDEFGNADILEATYPAVARLHGDW 120
 QY 260 QDPALPVVTGFLGKGMKSGAVTTLLRGSSDLTATTIGKALGLREIOWKVDVGLTCDPN 319
 Db 121 QDPALPVVTGFLGKGMKSGAVTTLLRGSSDLTATTIGKALGLREIOWKVDVGLTCDPN 180
 QY 320 IYPAKTYVYLTFEEATELAYFGAQLHPQSRPARBEDI PVKKSYNKAPGTLITRQ 379
 Db 181 IYPAKTYVYLTFEEATELAYFGAQLHPQSRPARBEDI PVKKSYNKAPGTLITRQ 240
 QY 380 RMDXGLVLTSLVLSKNTMLDIYSTMLGQYGLARVSGICYIEDLCISVDCVATSEV 439
 Db 241 RMDX--VLTSLVLSKNTMLDIYSTMLGQYGLARVFAI--FEDLCISVDCVATSEV 296

```

QY 440 SVSVSLDPSKINRELIIQASLHDVVEELEKAIYRLIQORAIISLIGNVQSSLILEK 499
D 297 SVSVSLDPSKINRELIIQ--ELDHVVEELEKAIYRLIQORAIISLIGNVQSSLILEK 354
QY 500 TGSVLKSGVNVQMISQASKVMMSLIYHDSDAKALVEALHOAFEDDVLISQVEANLIV 559
D 355 TGSVLKSGVNVQMISQASKVMMSLIYHDSDAKALVEALHOAFEDDVLISQVEANLIV 414
QY 560 G 560
D 415 G 415

```

RESULT 4

```

US-10-424-599-170907
; Sequence 170907, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David X
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285664
; SEQ ID NO 170907
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(564)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_125343C.1.Pep
US-10-424-599-170907

```

```

Query Match 65.4%; Score 1839; DB 12; Length 564;
Best Local Similarity 71.0%; Pred. No. 6,3e-172;
Matches 391; Conservative 40; Mismatches 80; Indels 40; Gaps 7;

```

```

QY 37 GPRGARGLSMVVADSNRRRAKQADGDVIGAPVL----- 71
D 10 GVQGLKLAIVMSVRSRSLHCKSQI--GFALGAPVCARRWGNRVAFSVTTCKASTDYIE 67
QY 72 -----GGLMGEGLDQSLVWVKFGSSSVSAARMAVAGLITPPEERPVVYLSAMGKT 125
D 68 KNAITENGMVSSG-ETSFCTVMKFGSSSVASADRMKEVATLILSPPEERPIVYLSAMGKT 126
QY 126 TNNLLAGKRAVGGCVIHSEIEEMNMVSLHIKIVDELGLP-XICNTSLVELQLLKI 184
D 127 TNNLLAGKRAVGGCVIHSEIEEMNMVSLHIKIVDELGLP-XICNTSLVELQLLKI 186
QY 185 AMMKELTPRTSDYLVSFGEQSTRIFSAVYLNKIRVAKQYDAFDIGFTTDFGNADILE 244
D 187 AMMKELTPRTSDYLVSFGEQSTRIFSAVYLNKIRVAKQYDAFDIGFTTDFGNADILE 246
QY 245 ATYPAVAKRLHGMWIDPAIPVYTGFLGKMGSGAVTTLGRGSDLTATTTIGKALGLREI 304
D 247 ATYPAVAKRLHGMWIDPAIPVYTGFLGKMGSGAVTTLGRGSDLTATTTIGKALGLREI 306
QY 305 QVMKDVGVLTCDPNITPFAKTPVYLTFEATLAFGAQVYLHPOSMRPAAREGDIPIRYK 364
D 307 QVMKDVGVLTCDPNITPFAKTPVYLTFEATLAFGAQVYLHPOSMRPAAREGDIPIRYK 366
QY 365 NSYNPAPGTLITRORDMXGLVLTSLVLSKNTVMDIVSTRMLGQYGFILARVSGICYI 424
D 367 NSYNPAPGTLITRORDMXGLVLTSLVLSKNTVMDIVSTRMLGQYGFILARVSGICYI 422
QY 425 EDLCISVDCAVTSVSVSLDPSKINRELIIQASLHDVVEELEKAIYRLIQORAI 484

```

```

D 423 EELGISVDVATSEVSVSLTLDPSKINRELIIQASLHDVVEELEKAIYRLIQORAI 482
QY 485 SLIGNVQSSLILEKTRGVLRKSGVNVQMISQASKVMMSLIYHDSDAKALVEALHOAF 544
D 483 SLIGNVQSSLILEKTRGVLRKSGVNVQMISQASKVMMSLIYHDSDAKALVEALHOAF 542
QY 545 EDDVLSQVEAE 555
D 543 ESE-LSLEMD 552

```

RESULT 5

```

US-09-890-813-17
; Sequence 17, Application US/09890813
; Publication No. US20020183486A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Aspartate Kinase
; FILE REFERENCE: BB1430 PCT
; CURRENT APPLICATION NUMBER: US/09/890,813
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/172944
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Glycine max
US-09-890-813-17

```

```

Query Match 65.2%; Score 1835; DB 9; Length 564;
Best Local Similarity 70.6%; Pred. No. 1.6e-171;
Matches 389; Conservative 42; Mismatches 80; Indels 40; Gaps 7;

```

```

QY 37 GPRGARGLSMVVADSNRRRAKQADGDVIGAPVL----- 71
D 10 GVQGLKLAIVMSVRSRSLHCKSQI--GFALGAPVCARRWGNRVAFSVTTCKASTDYIE 67
QY 72 -----GGLMGEGLDQSLVWVKFGSSSVSAARMAVAGLITPPEERPVVYLSAMGKT 125
D 68 KNAITENGMVSSG-ETSFCTVMKFGSSSVASADRMKEVATLILSPPEERPIVYLSAMGKT 126
QY 126 TNNLLAGKRAVGGCVIHSEIEEMNMVSLHIKIVDELGLP-XICNTSLVELQLLKI 184
D 127 TNNLLAGKRAVGGCVIHSEIEEMNMVSLHIKIVDELGLP-XICNTSLVELQLLKI 186
QY 185 AMMKELTPRTSDYLVSFGEQSTRIFSAVYLNKIRVAKQYDAFDIGFTTDFGNADILE 244
D 187 AMMKELTPRTSDYLVSFGEQSTRIFSAVYLNKIRVAKQYDAFDIGFTTDFGNADILE 246
QY 245 ATYPAVAKRLHGMWIDPAIPVYTGFLGKMGSGAVTTLGRGSDLTATTTIGKALGLREI 304
D 247 ATYPAVAKRLHGMWIDPAIPVYTGFLGKMGSGAVTTLGRGSDLTATTTIGKALGLREI 306
QY 305 QVMKDVGVLTCDPNITPFAKTPVYLTFEATLAFGAQVYLHPOSMRPAAREGDIPIRYK 364
D 307 QVMKDVGVLTCDPNITPFAKTPVYLTFEATLAFGAQVYLHPOSMRPAAREGDIPIRYK 366
QY 365 NSYNPAPGTLITRORDMXGLVLTSLVLSKNTVMDIVSTRMLGQYGFILARVSGICYI 424
D 367 NSYNPAPGTLITRORDMXGLVLTSLVLSKNTVMDIVSTRMLGQYGFILARVSGICYI 422
QY 425 EDLCISVDCAVTSVSVSLDPSKINRELIIQASLHDVVEELEKAIYRLIQORAI 484
D 423 EELGISVDVATSEVSVSLTLDPSKINRELIIQASLHDVVEELEKAIYRLIQORAI 482
QY 485 SLIGNVQSSLILEKTRGVLRKSGVNVQMISQASKVMMSLIYHDSDAKALVEALHOAF 544
D 483 SLIGNVQSSLILEKTRGVLRKSGVNVQMISQASKVMMSLIYHDSDAKALVEALHOAF 542
QY 545 EDDVLSQVEAE 555

```

Db 543 ESE-ISELEMD 552

```

RESULT 6
US-10-425-114-52733
; Sequence 52733, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 52733
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700556108_FLI.pep
US-10-425-114-52733

```

Query Match 64.9%; Score 1826.5; DB 12; Length 493;
 Best Local Similarity 77.7%; Pred. No. 8.7e-171;
 Matches 377; Conservative 39; Mismatches 62; Indels 7; Gaps 5;

```

QY 72 GCGMGBGLDQLSVNKKFGSSSVSAAAEVAGLITPEERPVVVSAMGKTNNLL 131
DB 3 GWSVSSG-ETSFVCKMFGSSSVASADRMKEVATLILSPBERPLVVSAMGKTNNLL 61
QY 132 AGEKAVGCGVHYHSEIEEMNWKSLIKITVDELGLP-XICNTSLYEELQIKIAMMEL 190
DB 62 AGEKAVGCGVHYHSEIEEMNWKSLIKITVDELGLP-XICNTSLYEELQIKIAMMEL 121
QY 191 TPRTSDYLVSPFGCMSTRIFSAYLNKIRKARQYADPDIGFTTPEFGNADILETAYAV 250
DB 122 TKTODYLVSPFGCMSTRIFSAYLNKIRKARQYADPDIGFTTPEFGNADILETAYAV 161
QY 251 AKLHGDMDIPALPVYTGFLGKMGSGAVTTLGRGSDLTATTIGKALGREIQWMDV 310
DB 182 AKLHGDMDIPALPVYTGFLGKMGSGAVTTLGRGSDLTATTIGKALGREIQWMDV 241
QY 311 DGVLTCDPNITPAKTVPPYITFEATELAYFGAQLVHPOSMPAREGDIPIRVKNSYNPK 370
DB 242 DGVLTCDPNITPAKTVPPYITFEATELAYFGAQLVHPOSMPAREGDIPIRVKNSYNPK 301
QY 371 APGTLTTRORDMXXGLVLTSLYKSNVTMLDIVSTRMGQYGFARVSGLCYIEDLCIS 430
DB 302 APGTLTTRORDMXXGLVLTSLYKSNVTMLDIVSTRMGQYGFARVSGLCYIEDLCIS 357
QY 431 VDCVATSEVSVALDPSKTIWRELLIQASELDHVEEIEKIALYVRLLOQRAIISLIINV 490
DB 358 VDCVATSEVSVALDPSKTIWRELLIQASELDHVEEIEKIALYVRLLOQRAIISLIINV 417
QY 491 EGSLLIEKTRGVLRKSGVNVQMIISQASKYNMSTLVHSDAKALVEALHQAFFEDVLS 550
DB 418 EGSLLIEKTRGVLRKSGVNVQMIISQASKYNMSTLVHSDAKALVEALHQAFFEDVLS 476
QY 551 QVEAE 555
DB 477 ELEMND 481

```

RESULT 7
 US-10-424-599-170902
 ; Sequence 170902, Application US/10424559

```

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5323)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 170902
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_125339C.1.pep
US-10-424-599-170902

```

Query Match 62.4%; Score 1754.5; DB 12; Length 561;
 Best Local Similarity 71.9%; Pred. No. 1.3e-163;
 Matches 361; Conservative 52; Mismatches 74; Indels 15; Gaps 6;

```

QY 55 RAKQADGGDGVLAGAVLGLGMBGLDQLSVNKKFGSSSVSAAAEVAGLITPEER 114
DB 62 KATESDVEGEGS-----GFATEMS YTCVMKFGSSSVANERKEVANLILSPBER 114
QY 115 PVVLSAMGKTNNLLAGEKAVGCGVHYHSEIEEMNWKSLIKITVDELGLP-XICNTS 173
DB 115 PVVLSAMGKTNNLLAGEKAVGCGVHYHSEIEEMNWKSLIKITVDELGLP-XICNTS 174
QY 174 LYEELQIKIAMMELTPRTSDYLVSPFGCMSTRIFSAYLNKIRKARQYADPDIGFTT 233
DB 175 LYEELQIKIAMMELTPRTSDYLVSPFGCMSTRIFSAYLNKIRKARQYADPDIGFTT 234
QY 234 TPEFGNADILETAYAVAKRLHGDMDIPALPVYTGFLGKMGSGAVTTLGRGSDLTAT 293
DB 235 TPEFGNADILETAYAVAKRLHGDMDIPALPVYTGFLGKMGSGAVTTLGRGSDLTAT 294
QY 294 TIGKALGREIQWMDVGVLTCDPNITPAKTVPPYITFEATELAYFGAQLVHPOSMP 353
DB 295 TIGKALGREIQWMDVGVLTCDPNITPAKTVPPYITFEATELAYFGAQLVHPOSMP 354
QY 354 AREGDIPIRVKNSYNPKAPGTLITRQPMXGLVLTSLYKSNVTMLDIVSTRMGQY 413
DB 355 AREGDIPIRVKNSYNPKAPGTLITRQPMXGLVLTSLYKSNVTMLDIVSTRMGQY 412
QY 414 FLARVSGLCYIEDLCISVDCVATSEVSVALDPSKTIWRELLIQASELDHVEEIEKIA 473
DB 413 FLARVSGLCYIEDLCISVDCVATSEVSVALDPSKTIWRELLIQASELDHVEEIEKIA 468
QY 474 IYRLLOQRAIISLIINVQSSILIEKTRGVLRKSGVNVQMIISQASKYNMSTLVHSD 533
DB 469 IYRLLOQRAIISLIINVQSSILIEKTRGVLRKSGVNVQMIISQASKYNMSTLVHSD 528
QY 534 ACVEALHQAFFEDVLSQVEAE 555
DB 529 QCVRALHKAFFESE-ISELEND 549

```

RESULT 8
 US-09-890-813-16
 ; Sequence 16, Application US/09890813
 ; Publication No. US20020183486A1
 ; GENERAL INFORMATION:
 ; APPLICANT: E.I. du Pont de Nemours and Company
 ; TITLE OF INVENTION: Aspartate Kinase
 ; FILE REFERENCE: B81430 PCT
 ; CURRENT APPLICATION NUMBER: US/09/890,813
 ; CURRENT FILING DATE: 2001-08-02
 ; PRIOR APPLICATION NUMBER: 60/172944
 ; PRIOR FILING DATE: 1999-12-21

NUMBER OF SEQ ID NOS: 24
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 16
 LENGTH: 439
 TYPE: PRT
 ORGANISM: *Triticum aestivum*
 US-09-890-813-16

Query Match 57.9%; Score 1629.5; DB 9; Length 439;
 Best Local Similarity 77.7%; Pred. No. 1.9e-151;
 Matches 331; Conservative 37; Mismatches 51; Indels 7; Gaps 4;

QY 133 GEAAGCGVYHSEIEEMNNVKSILHKTVDLGLP-XICNTSLYELEQLKGIAMMKELTPRTSDYLSVSGE 191
 DB 2 GTRAVSCGAPKASEIYELAVIKELHRTIDELGSSIVSGFLDEQLKGVAMMKELT 61
 QY 192 PRSDYLSVSGECMSTRISAYLNKIRVAKROYDADFIGITTFDEFGNADILEATYPAPA 251
 DB 62 LRTRDYLSVSGECMSTRISAYLNKIRVAKROYDADFIGITTFDEFGNADILEATYPAPA 121
 QY 252 KRLHGMIDDPALPVYTGFLGKMGSGAVTTLGRGSDLTATTIGALGLREIQWKD 311
 DB 122 KRLHGMIDDPALPVYTGFLGKMGSCAVTTLGRGSDLTATTIGALGLREIQWKD 181
 QY 312 GVLTCDPNTYPAKKTVPYLTPEEATLAYGAQVLAHPOSMPAREGDIPIRVKNSYNPKA 371
 DB 182 GVLTCDPNTYPAKKTVPYLTPEEATLAYGAQVLAHPOSMPAREGDIPIRVKNSYNPKA 241
 QY 372 PGTLITRORDMDXGLVYLTSLVLSKSNVTMLDIVSTRMLGQYGLFARVSGICYIEDLCISV 431
 DB 242 PGTVITKTRDMRKS--ILTSIVLKSNTIMLDIVSTRMLGQYGLFARVSGI--FEDLGISV 297
 QY 432 DCVATSEVSVSLDSSKIMSRELIQAASELDHVEELEKIAIVRLQORATISLIGNVE 491
 DB 298 DSVATSEVSVSLDSSKIMSRELIQA--ELDHVEELEKIAIVRLQORATISLIGNVQ 355
 QY 492 QSSLLIEKTGRVLRKSGVNVQMSIQGASKVNSLIYHSDAKALVEALHQAFFEDVLSQ 551
 DB 356 RSSLLIEKAFNVLRKSGVNVQMSIQGASKVNSLIYHSDAKALVEALHQAFFEDVLSQ 415
 QY 552 VEAENL 557
 DB 416 VEEADL 421

RESULT 9

US-10-425-114-70672
 Sequence 70672, Application US/10425114
 Publication No. US20040034888A1
 GENERAL INFORMATION:
 APPLICANT: Liu, Jindong
 APPLICANT: Zhou, Yihua
 APPLICANT: Kovalic, Steven E
 APPLICANT: Screen, Steven E
 APPLICANT: Tabaska, Jack E
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 FILE REFERENCE: 38-21(5313)B
 CURRENT APPLICATION NUMBER: US/10/425,114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 70672
 LENGTH: 426
 TYPE: PRT
 ORGANISM: *Zea mays*
 FEATURE:
 OTHER INFORMATION: Clone ID: 700343285_FLI.pep
 US-10-425-114-70672

Query Match 56.4%; Score 1587.5; DB 12; Length 426;
 Best Local Similarity 78.8%; Pred. No. 2.5e-147;
 Matches 323; Conservative 37; Mismatches 43; Indels 7; Gaps 4;

QY 145 SEIEEMNNVKSILHKTVDLGLP-XICNTSLYELEQLKGIAMMKELTPRTSDYLSVSGE 203
 DB 3 SEIPELAVIKDILHRTVDLGLDRSIVSGFLDEQLKGVAMMKELTPRTDYLSVSGE 62
 QY 204 CMSTRISAYLNKIRVAKROYDADFIGITTFDEFGNADILEATYPAPAKLHGMIDPA 263
 DB 63 CMSTRISAYLNKIRVAKROYDADFIGITTFDEFGNADILEATYPAPAKLHGMIDPA 122
 QY 264 IYVYTGFLGKMGSGAVTTLGRGSDLTATTIGALGLREIQWKDVGVLTCDPNTYPA 323
 DB 123 IYVYTGFLGKMGSCAVTTLGRGSDLTATTIGALGLREIQWKDVGVLTCDPNTYPA 182
 QY 324 ACTVYLTPEEATLAYGAQVLAHPOSMPAREGDIPIRVKNSYNPKAPGTLITRORDMD 383
 DB 183 ALPVPYLTPEEATLAYGAQVLAHPOSMPAREGDIPIRVKNSYNPKAPGTLITRORDMD 242
 QY 384 XGLVYLTSLVLSKSNVTMLDIVSTRMLGQYGLFARVSGICYIEDLCISVDCVATSEVSV 443
 DB 243 KS--ILTSIVLKSNTIMLDIVSTRMLGQYGLFARVSGI--FEDLGISVDCVATSEVSV 298
 QY 444 SLDPSKIMSRELIQAASELDHVEELEKIAIVRLQORATISLIGNVQSSLLIEKAFNV 503
 DB 299 TLDPSKIMSRELIQA--ELDHVEELEKIAIVRLQORATISLIGNVQSSLLIEKAFNV 356
 QY 504 LRKSGVNVQMSIQGASKVNSLIYHSDAKALVEALHQAFFEDVLSQVE 553
 DB 357 LRKSGVNVQMSIQGASKVNSLIYHSDAKALVEALHQAFFEDVLSQVE 406

RESULT 10

US-10-425-114-39779
 Sequence 39779, Application US/10425114
 Publication No. US20040034888A1
 GENERAL INFORMATION:
 APPLICANT: Liu, Jindong
 APPLICANT: Zhou, Yihua
 APPLICANT: Kovalic, Steven E
 APPLICANT: Screen, Steven E
 APPLICANT: Tabaska, Jack E
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 FILE REFERENCE: 38-21(5313)B
 CURRENT APPLICATION NUMBER: US/10/425,114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 39779
 LENGTH: 426
 TYPE: PRT
 ORGANISM: *Glycine max*
 FEATURE:
 OTHER INFORMATION: Clone ID: 700747979_FLI.pep
 US-10-425-114-39779

Query Match 55.9%; Score 1573; DB 12; Length 426;
 Best Local Similarity 77.7%; Pred. No. 6.8e-146;
 Matches 324; Conservative 34; Mismatches 51; Indels 8; Gaps 5;

QY 140 GVLHVEIEEMNNVKSILHKTVDLGLP-XICNTSLYELEQLKGIAMMKELTPRTSDYL 198
 DB 1 GVTNVSIEELCPFKDILHRTVDLGLDGSVIAKLEIEQLKGIAMMKELTKRTQY 60
 QY 199 VSFGECCMSTRISAYLNKIRVAKROYDADFIGITTFDEFGNADILEATYPAPAKLHGM 258
 DB 61 VSFGECCMSTRISAYLNKIRVAKROYDADFIGITTFDEFGNADILEATYPAPAKLHGM 120
 QY 259 IODPALPVYTGFLGKMGSGAVTTLGRGSDLTATTIGALGLREIQWKDVGVLTCDP 318
 DB 121 IODPALPVYTGFLGKMGSCAVTTLGRGSDLTATTIGALGLREIQWKDVGVLTCDP 180
 QY 319 NIYPHAKTVYLTPEEATLAYGAQVLAHPOSMPAREGDIPIRVKNSYNPKAPGTLITR 378

Db 181 NIYPKAEPPVYLTFDEAALAFGAQVLAHPQSMRPAESDIPVRKNSYNPKAPGTLITK 240
 QY 379 QRDMDXGLVLTISIVKSNVTMLDIYSTRLMGQYFPLARVSGICIEDLCTISVDCVATSE 438
 Db 241 ARDMSK--AVLTSTIVKRVNVTMLDIYSTRLMGQYFPLAKVFSEI--FEEELGISVAVATSE 296
 QY 439 VASVSLDPSKIMSRRELIOQASLHDVVELEKIAVLLQORATISLIGNVEOSSLILE 498
 Db 297 VASVSLDPSKIMSRRELIOQ--ELDHVVELEKIAVNLQORATISLIGNVQSSLILE 354
 QY 499 KTGRLRKSGVNVQMSQASQKYNLSLVHSDAVALYEAHQAFEDDVLISQVEAE 555
 Db 355 KARRVLRITGITVQMSQASQKYNLSLVNVDSEACVRAHLAFPESE--LSELEMD 410

RESULT 11
 US-10-425-114-39716
 ; Sequence 39716; Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 39716
 ; LENGTH: 326
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700663367_FLI.pep
 ; US-10-425-114-39716

Query Match 43.5%; Score 1222.5; DB 12; Length 326;
 Best Local Similarity 79.5%; Pred. No. 1.6e-111;
 Matches 252; Conservative 24; Mismatches 34; Indels 7; Gaps 4;

QY 239 NADILEATYPAVAKRLHGMIDOPALPVYTGFLGKMKSGAVTTTIGRGSDLTATTICKA 298
 Db 1 NADILEATYPAVAKRLHGMIDOPALPVYTGFLGKMKSCAVTTTIGRGSDLTATTICKA 60
 QY 299 LGRREIQWKCVDVGLTCDPNIYPHAKTVPYLTFEATELAFYGAQVLAHPQSMRPAESGD 358
 Db 61 LGRREIQWKCVDVGLTCDPNIYPHAKTVPYLTFEATELAFYGAQVLAHPQSMRPAESGD 120
 QY 359 IPVRVKNKSNPKAPGTLITRQMDXGLVLTISIVKSNVTMLDIYSTRLMGQYFPLARV 418
 Db 121 IPVRVKNKSNPKAPGTLITRQMDXGLVLTISIVKSNVTMLDIYSTRLMGQYFPLARV 178
 QY 419 SGCIYIEDLCTISVDCVATSEVSVSLDPSKIMSRRELIOQASLHDVVELEKIAVLL 478
 Db 179 FSEI--FEEELGISVAVATSEVSVSLDPSKIMSRRELIOQ--ELDHVVELEKIAVNL 234
 QY 479 QORATISLIGNVEOSSLILEKTRGVLRKSGVNVQMSQASQKYNLSLVHSDAVALYEA 538
 Db 235 QORATISLIGNVQSSLILEKTRGVLRKSGVNVQMSQASQKYNLSLVNVDSEACVRAH 294
 QY 539 LHOAFEDDVLISQVEAE 555
 Db 295 LHLAFPESE--LSELEMD 410

RESULT 12
 US-10-425-114-43083
 ; Sequence 43083; Application US/10425114
 ; Publication No. US20040034888A1

; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 43083
 ; LENGTH: 249
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700664751_FLI.pep
 ; US-10-425-114-43083

Query Match 35.3%; Score 992.5; DB 12; Length 249;
 Best Local Similarity 78.1%; Pred. No. 5e-89;
 Matches 193; Conservative 21; Mismatches 32; Indels 1; Gaps 1;

QY 99 RMAEYAGLITFPERPVPVYLSANGKTTNNLLGGERAVGCGVTHVSEIEMNMYKSLHT 158
 Db 2 RREYVANNILTSFERPPIIVSANGKTTNNLLGGERAVGCGVTHVSEIEMNMYKSLHT 61
 QY 159 KTVDELGLP-XICNTSLYELBOLKGIAMKELTPRTSDYVSFGECMSTRISAYLNKI 217
 Db 62 RTVEQLGVDRNVIEHLELQKLGIMKELTPRODYVVSFGECMSTRISAYLNKI 121
 QY 218 RYKARQYDAFDIGFTTDFEGNADILEATYPAVAKRLHGMIDOPALPVYTGFLGKMK 277
 Db 122 GYKARQYDAFEMGIITTDFTNADILEATYPAVAKRLHSDVSPDPAIPVYTGFLGKMK 181
 QY 278 GAVTTIGRGSDLTATTICKALGRIQWKCVDVGLTCDPNIYPHAKTVPYLTFEATE 337
 Db 182 CAVTTIGRGSDLTATTICKALGRIQWKCVDVGLTCDPNICPQAKPPYLTDEAE 241
 QY 338 LAYFGAQ 344
 Db 242 LAYFGAQ 248

RESULT 13
 US-10-425-114-54453
 ; Sequence 54453; Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 54453
 ; LENGTH: 221
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: UC-ZMFLM017134C03_FLI.pep
 ; US-10-425-114-54453

Query Match 34.6%; Score 972; DB 12; Length 221;
 Best Local Similarity 95.4%; Pred. No. 4.3e-87;

```

Matches 207; Conservative 0; Mismatches 4; Indels 6; Gaps 3;
QY 344 QVLPQSMRPARBEDIPIRVKNSYNPKAPGTLLITRQDMXGLVLTSLVLSXNTYMDI 403
Db 11 QVLPQSMRPARBEDIPIRVKNSYNPKAPGTLLITRQDMXGLVLTSLVLSXNTYMDI 68
QY 404 VSTRMFGVGFILARVSGICIEELCTSVDCVATSEVSISLDPKSKMRELIQASELD 463
Db 69 VSTRMFGVGFILARVSGICIEELCTSVDCVATSEVSISLDPKSKMRELIQASELD 124
QY 464 HVEBELEKIAIVRLLOQRAIISLIGVNEGSSILLEKTGRVLRKSGVNVQMISQASKVM 523
Db 125 HVEBELEKIAIVRLLOQRAIISLIGVNEGSSILLEKTGRVLRKSGVNVQMISQASKVM 184
QY 524 SLIVHSDAKALVLAHQAFEDDVLSQVEAENLVG 560
Db 185 SLIVHSDAKALVLAHQAFEDDVLSQVEAENLVG 221

```

RESULT 14

```

US-09-890-813-12
; Sequence 12, Application US/09890813
; Publication No. US20020183486A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Aspartate Kinase
; FILE REFERENCE: BRL430 PCT
; CURRENT APPLICATION NUMBER: US/09/890,813
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/172944
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-890-813-12

```

```

Query Match 31.7%; Score 890.5; DB 9; Length 281;
Best Local Similarity 67.1%; Pred. No. 6.9e-79;
Matches 192; Conservative 19; Mismatches 68; Indels 7; Gaps 4;

```

```

QY 19 PPASSGHVRLAC-FGRTTGPGRAGLSMNVYADSTRRAKQADGSGVLCAPVLSGLME 77
Db 1 PPRVGRBQVYLAACAARPGRCRRRGLVRCQSGAAAVLNKDD--AAAYAAAAAASS 57
QY 78 GLDQLSVYMKFGGSSVSAARMVAAGLITFPEERPVVLSAMGKTNNLLAGEKAV 137
Db 58 ATG--FTYAMKFGGSSVSAARMREVAADLISPEETPVVLSAMGKTNNLLAGEKAV 115
QY 138 GCGVIVHSEIEEMNVKSLHIKIVDELGLP-XICNTSLVLEQLKGIAMKELTPRISD 196
Db 116 SCGAPKASEIPELAVIKELHRTIDELGLDRSIVSGLEBEQLKGVAMKELTPRISD 175
QY 197 YLVFSGECMSTRISAVYLNKIRVAKROYDAFDIGFITTFDEFGNADILEATYPVAKRLHG 256
Db 176 YLVFSGECMSTRIPAYLNKIGKARQYDAFDIGFITTFDFTNADILEATYPVAKRLHG 235
QY 257 DWIDDPALPVVTGFLGKMGSGAVTTLGRGSDLTATTIGKALGLR 302
Db 226 DWIDDPALPVVTGFLGKMGSCAVTTLGRGSDLTATTIGKALRTR 281

```

RESULT 15

```

US-10-425-114-46004
; Sequence 46004, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E

```

```

; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 72128
; SEQ ID NO 46004
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700873216_FLI.pep
US-10-425-114-46004

```

```

Query Match 27.1%; Score 762.5; DB 12; Length 226;
Best Local Similarity 73.9%; Pred. No. 2e-66;
Matches 164; Conservative 22; Mismatches 27; Indels 9; Gaps 4;

```

```

QY 335 ATELAVFGAQLHPQSMRPARBEDIPIRVKNSYNPKAPGTLLITRQDMXGLVLTSLV 394
Db 1 ATELAVFGAQLHPQSMRPARBEDIPIRVKNSYNPKAPGTLLITRQDMXGLVLTSLV 58
QY 395 KSNVTMLDIVSTMLOQYGFILARVSGICIEELCTSVDCVATSEVSISLDPKSKMRE 454
Db 59 KSNVTMLDIVSTMLOQYGFILARVSGICIEELCTSVDCVATSEVSISLDPKSKMRE 116
QY 455 LIQASELDHVELEKIAIVRLLOQRAIISLIGVNEGSSILLEKTGRVLRKSGVNVQM 514
Db 117 LIQ--ELDVVELEKIAIVRLLOQRAIISLIGVNEGSSILLEKTGRVLRKSGVNVQM 174
QY 515 SQGASKVMISLIVHSDAKALVLAHQAFEDDVLSQVEAENLVG 556
Db 175 SQGASKVMISLIVHSDAKALVLAHQAFEDDVLSQVEAENLVG 213

```

```

Search completed: March 23, 2004, 13:23:17
Job time : 68 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 23, 2004, 13:02:31 ; Search time 27 Seconds
(without alignments)
1995.084 Million cell updates/sec

Title: US-09-890-813-6
Perfect score: 2813
Sequence: 1 MAIPVRSAAAPRLVPSIPP.....QAFPEDVLSQVAENLLVG 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1789	63.6	569	2 T48575	aspartate kinase
2	662	23.5	473	2 T48371	aspartate kinase
3	642.5	22.8	462	2 D69337	aspartate kinase
4	607	21.6	473	2 P75405	aspartate kinase
5	590	21.0	461	2 AG0452	aspartate kinase
6	569	20.2	449	1 KIECD3	aspartate kinase
7	569	20.2	449	2 C86095	aspartate kinase
8	569	20.2	449	2 G91254	lysine sensitive a
9	566	20.1	449	2 AC1013	aspartate kinase
10	531	18.9	479	2 P82328	aspartate kinase
11	516.5	18.4	519	2 T39822	aspartate kinase
12	491.5	17.5	527	1 KIRBD	aspartate kinase
13	483.5	17.2	825	2 H82086	aspartate kinase
14	477.5	17.0	820	1 DEBCK	thra difunctional
15	477.5	17.0	820	2 B85480	aspartate kinase
16	477.5	17.0	820	2 B90629	aspartate kinase
17	475.5	16.9	815	2 A64048	aspartate kinase
18	469.5	16.7	820	2 AC0502	aspartate kinase
19	468.5	16.7	820	2 F97918	aspartate kinase
20	467.5	16.6	454	2 G95047	aspartate kinase
21	464.5	16.5	828	2 G82083	aspartate kinase
22	463.5	16.5	819	2 B47057	aspartate kinase
23	461.5	16.4	916	2 T02952	aspartate kinase
24	461.5	16.4	917	2 T02954	aspartate kinase
25	455.5	16.2	819	2 AH0056	aspartate kinase
26	452.5	16.1	920	2 T02953	aspartate kinase
27	451	16.0	911	2 S46497	aspartate kinase
28	447	15.9	911	2 E86438	hypothetical prote
29	445	15.8	437	2 H96933	aspartate kinase

30	442	15.7	473	2 E72715	probable aspartate
31	441	15.7	739	2 H72364	aspartate kinase II -
32	439	15.6	915	2 T03589	probable aspartate
33	429	15.3	446	2 D90238	aspartate kinase (AKH
34	425.5	15.1	460	2 H89907	hypothetical prote
35	425	15.1	868	2 A82722	difunctional diaml
36	423.5	15.0	916	2 T06242	aspartate kinase
37	420.5	14.9	909	2 T06246	aspartate kinase
38	417.5	14.8	421	2 S15276	aspartate kinase
39	416.5	14.8	456	2 D83837	homoserine dehydro
40	415.5	14.8	421	2 S42422	aspartate kinase
41	412.5	14.7	921	2 S35160	aspartate kinase
42	409.5	14.6	421	2 F70794	probable ask prote
43	408.5	14.5	454	2 A69763	homoserine dehydro
44	408	14.5	399	2 D97123	aspartate kinase
45	402	14.3	411	2 A48946	aspartate kinase

ALIGNMENTS

RESULT 1

T48575
N/Alternate names: Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C/Accession: T48575
R/Reven, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancro
Submitted to the Protein Sequence Database, April 2000
A/Reference number: Z2490
A/Accession: T48575
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-569 <BEV>
A/Cross-references: EMBL:AL163491
A/Experimental source: cultivar Columbia; BAC clone T31B5
C/genetics:
A/Map position: 5
A/Introns: 134/3; 162/2; 176/1; 227/3; 277/3; 309/3; 348/3; 388/3; 419/3; 458/3; 497/
A/Note: T31B5.100

Query Match	Best Local Similarity	Score	DB 2	Length	569
Matches	369	71.8%	54	Mismatches	79
				Indels	12
				Gaps	6
QY	44	LSMVVADSTSRRAKQADGDVLAGPVILGMEGLD---	QISVVMKFGSSVSAAAR	99	
DB	44	LSLPIDGDSIRKVGSGSRNIVRA-VLEKTEATEVDEKGITVCKFGSSVSAER	102		
QY	100	MAEVAQLITPEEPFVVLISAMGKTNNLLAEKXVGVHVSIEEMNMYKSLHIK	159		
DB	103	MEEVADLLITPEESPVLISAMGKTNNLLAEKXVGVHVSIEEMNMYKSLHIK	162		
QY	160	TYDEIGL-FLICNTSLVELEQLKGIAMKELTPRTDYLVSFECMSTRIFSAVYLNIR	218		
DB	163	TYKELNIDPSVILTYLEELQELKGIAMKELTRIDYVSPFECSTRIFSAVYLNITIG	222		
QY	219	VVAROYDAFDITTTDEFGADILEAVYPAVAKRLGDMVQDPAIVVMGFLGKMGKSG	278		
DB	223	VVAROYDAEIGFITTDFTMGDILEAVYPAVAKRLGDMVQDPAIVVMGFLGKMGKSG	282		
QY	279	AVTTLGRGSDLTATTIGALGLREIQWKKVDVGLTCDPNYIFHAKTVPYLFEAEATP	336		
DB	283	AVTTLGRGSDLTATTIGALGLREIQWKKVDVGLTCDPNYIFHAKTVPYLFEAEATP	342		
QY	339	AIFGAQVLPQSMRPARBEDIIVRYKNSYNPKAPETLITRQMDXGIVLTSLVKSANV	398		
DB	343	AIFGAQVLPQSMRPARBEDIIVRYKNSYNPKAPETLITRQMDXGIVLTSLVKSANV	400		
QY	399	TMDIYSTRLGQYGLIARVSGICVIEDICISVDVATSEVSVSLDPSKIMSRRELIOQ	458		
DB	401	TMDIYSTRLGQYGLIARVSGICVIEDICISVDVATSEVSVSLDPSKIMSRRELIOQ	458		

QY 459 ASELDHVEELEKIAIVRLLOQRAIIISLIGNVEOSSILIEKTRGVLRKSGVNVOMISQGA 518
 DB 459 --ELDHVVEELEKIAIVRLLOQRAIIISLIGNVEOSSILIEKTRGVLRKSGVNVOMISQGA 516
 QY 519 SKVMSLIVHSDKALVEALHQAFFEDVLSQV 552
 DB 517 SKVMSLIVHSDKALVEALHQAFFEDVLSQV 550
 RESULT 2
 C64371
 aspartate kinase (BC 2.7.2.4) I - Methanococcus jannaschii
 N:Alternate names: aspartokinase I
 C:Species: Methanococcus jannaschii
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 25-Aug-2003
 C:Accession: C64371
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
 Reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glodek, A.;
 Iron, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; MUID:96337999; PMID:8688087
 A:Accession: C64371
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-473 <BLU>
 A:Cross-references: GB:U67506; GB:L77117; NID:g1591274; PIDN:AA898565.1; PID:g1591278; T
 C:Genetics:
 A:Map position: FOR508401-509822
 C:Superfamily: aspartate kinase; aspartate kinase homology
 C:Keywords: phosphotransferase
 F:2-468/Domain: aspartate kinase homology <DKI>
 Query Match 23.5%; Score 662; DB 2; Length 473;
 Best Local Similarity 32.3%; Pred. No. 7.2e-39;
 Matches 160; Conservative 110; Mismatches 167; Indels 58; Gaps 11;
 QY 83 LSVWVKFGSSSVSSAARMAEYAGLITPEERP--VVVLSAMGKTNNLLAGEKAVGCG 140
 DB 1 MTTWVKFGSTVSGSERIRHVAKIYTKKKEDDDVVVVSAMSEVTALVEISQAL--D 58
 QY 141 VHVSEIEEM-NMVSILHKTVDL-----GLPKICNTSLYEELQLKGMKKEIT 191
 DB 59 VVDIAKVDGFIKPIREKHYKAIIEKIKSEETKEEKKIDSRILEKVLIGVAVGELT 118
 QY 192 PRTSDYIVSFGECMSTRIFSAVYLNKIRYKAROYDAFIDGITTDFGNA-----DILEAT 246
 DB 119 PKSRRIYLSFGRRLSSPLISGAIKRLGEXS:ALGGEAGGITDNNFGSAVVKLEVERL 178
 QY 247 YPAVAKRLHGMWIDPAPLIVYTGPIGKMKSGAVTTIGRGSDLTATTIGKALGREIQY 306
 DB 179 LPL-----LKEGIIIPVYTGFIGT-TEEGYITTLGRGSDVSALIGADADIT 228
 QY 307 KCDVQGVLTCDPNIYPAHAKTVPYLTPEBATELAYFGAQLHPQSMRPARBEDIPIRVKNS 366
 DB 229 WTDVSGVYTTDRLVPTARIPKISYIEMELAFGAKVHLPRITBPMKGIPIVKN 288
 QY 367 YNPKAPGTLITRQDMXGLVVLISVLSKSVNTMLDIVSTMYLQGYFLARVSGICTED 426
 DB 289 FEPESEGTLLITNDMSDSIVAKISTI--KNVALINIFGAGMVGSGTAARIFKALGEE 346
 QY 427 LGISDVCAVTSVSVSLDPSKIMSRRLIOASELDHVEELEK-----471
 DB 347 VAVVILISQGSSEITNIS-----LVVSEEDVAKALKALRREGDGRGKSLNNL 394
 QY 472 IAIIVRLLOQRAIIISLIG-NVEOSSILIEKTRGVLRKSGVNVOMISQGA SKVMSLIVHSD 530
 DB 395 IHDVSDKDVCAIVSYGAGMGAKIGACKIFTAVSESGANTKMTAOSSEVINISFVIDEK 454
 QY 531 DAKALVEALHQAFFE 545

DB 455 DLNCKVRKLEKFE 469
 RESULT 3
 D69337
 aspartate kinase (Iyec) homolog - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 25-Aug-2003
 C:Accession: D69337
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodge,
 Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Uitterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes,
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
 A:Reference number: A69250; MUID:96049343; PMID:9389475
 A:Accession: D69337
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-462 <KLE>
 A:Cross-references: GB:AE001056; GB:AE000782; NID:9268379; PIDN:AA890539.1; PID:9264;
 C:Superfamily: aspartate kinase; aspartate kinase homology
 F:2-459/Domain: aspartate kinase homology <DKI>
 Query Match 22.8%; Score 642.5; DB 2; Length 462;
 Best Local Similarity 34.4%; Pred. No. 1.6e-37;
 Matches 165; Conservative 88; Mismatches 187; Indels 39; Gaps 10;
 QY 83 LSVWVKFGSSSVSSAARMAEYAGLITPEER-ERPVVVLSAMGKTNNLLAGEKAV--G 138
 DB 1 MTTWVKFGSVSDGENTYHCANLVKFSNGENVVVVSAMQGVTDALLKAKRCCSESS 60
 QY 139 CGVI-----HVSEIF--EMNVSLHKTIVDELGLPKICNTSLYEELQLKGI 185
 DB 61 AGFIKMFIA DMKRYHAEIYAVRNDKIKAVIAVERL-----LDELEKYLIGIS 111
 QY 186 MKKELTRTDYIVSFGECMSTRIFSAVYLNKIRYKAROYDAFIDGITTDFGNDIIEA 245
 DB 112 YLGEITRSEDYIVSFEERLAPISFALISLIVSDVALTGDAGITTRNGRAKELPG 171
 QY 246 TVPAVAKRLHGMWIDPAPLIVYTGPIGKMKSGAVTTIGRGSDLTATTIGKALGREIQ 305
 DB 172 VYTTIRNRLEPLITKTIIPVYTGFIG-ATDDGSITTLGRGSDVYATILAAALDDEVW 230
 QY 306 VMKQVQGVLTCDPNIYPAHAKTVPYLTPEBATELAYFGAQLHPQSMRPARBEDIPIRVKNS 365
 DB 221 LMKEDVDMTCDDPKYVNAVIEPISYQEMELSHGAKILHPRALEPVRKNIPIRVKNS 290
 QY 366 SYNPKAPGTLITRQDMXGLVVLISVLSKSVNTMLDIVSTMYLQGYFLARVSGICTYE 425
 DB 231 TENDAGCTVIGPTTKOSSSEIVKALSLIPAGIVN--VSGAGFDPAELMSEVFGRLAE 347
 QY 426 DLTISDVCAVTSVSVSLDPSKIMSRRLIOASELDHVEELEKIAIVRLLOQRAIS 485
 DB 348 RVNVIYMAOSSSEINISLIVDIRDL--EKAYNALKELENGVIEKVIPIPI-----AVVS 399
 QY 486 LIGN-VREGSSILIEKTRGVLRKSGVNVOMISQGA SKVMSLIVHSDKALVEALHQAFF 543
 DB 400 AVSAGMAGTPEVAKITFSAIGKGINVIMISQGSSEVINISFVYDSDGYNKAVYHDEF 458
 RESULT 4
 F75405
 aspartate kinase - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 25-Aug-2003
 C:Accession: F75405
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.C.;
 M.; Shen, M.; Vamathevan, J.C.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: F75405
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-473 <WHI>
 A:Cross-references: GB:AE001982; GB:AE000513; NID:g6459109; PIDN:AAFI0936.1; PID:g64591.1
 A:Experimental source: strain R1
 A:Genetics:
 A:Gene: DK1365
 A:Map position:
 C:Superfamily: aspartate kinase; aspartate kinase homology

Query Match 21.6%; Score 607; DB 2; Length 473;
 Best Local Similarity 33.8%; Pred. No. 5,4e-35;
 Matches 163; Conservative 96; Mismatches 185; Indels 38; Gaps 14;

```

QY 85 VYMKFGSSVSSAARMAEVAAGLITFPEE--RPVVVLSAMGKTNNLLAGEKAVGCGVI 142
DB 11 LVMKFGGTVMQGAHAIRHSASLAGRSVEGVVVSAMAGVINGLHLADAAQSGDIA 70
QY 143 HVSEIEEMVMKSLHKTVDGLPXICNT-----SLYE-LEQLLKGIAMKELTPRTSD 196
DB 71 RAN--DEIAAKTRFETAQELGAAPDSETVRELHEETLRQAVGYVILLRETPRSR 128
QY 197 YVVFGEQWSTRIFSAYLNKIRVAKROYDADIGFITTTDEFQNDILEATVPAYAKRLG 256
DB 129 LVAFGERLSAPLMISALHQSGLRARHLTGEGAGILDTNFGNARPLPGTYERVGDRLSG 188
QY 257 DNIQDPAIPVVTGFLGKMKSGAVTTLGRGSDLTATTIGALGREIQWKDGVLTLC 316
DB 189 -FLSAGVTVIAGFME--TEGATITLGGGDTFSATTVGALGDEVMANKDVGWMSA 246
QY 317 DENIYFHAKEVYLTFEEATELAYGAQVLPQSRPARBEDIPIRVKNSYNKAPSTLI 376
DB 247 DRRVGDANINIVLSYGEVMEIAYGAKVLPALVATPLQDSGILPIRVKSADPPFAGTLV 306
QY 377 TQQRMDXGLVLTSLVKSNTMDIVSTRMLQGYGLARVSGICYIEDLCISVDCVAT 436
DB 307 RQQAEEIPRPVPAVATAK-NVSLITVAGAVLPEVIAVSFTALARENTTLIMVSQSS 365
QY 437 SEVSVSIDSPKINSRELIOQA-----SELDHVEELKEIAIYRLAQPAISLIGN- 489
DB 366 SMSNVSILAV-PSADARAL--QALRAGLSSELN--VEETNGV-----AVLAIGSG 411
QY 490 VQSSILIEKTRGVLRKSGVNTQMSQASKYNMILIYHSDPAALVEALHQAFFEDVL 549
DB 412 MKGQGVASARLPTALASEDINITIMTSGSSSELNISVALSGEVDPAVAHAFAF---TL 467
QY 550 SQ 551
DB 468 SQ 469
  
```

RESULT 5
 AG0452
 aspartate kinase (EC 2.7.2.4) [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 25-Aug-2003
 C:Accession: AG0452
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett, Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11566360
 A:Accession: AG0452
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-461 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC93187.1; PID:g15981636; GSPDB:GN00175
 C:Genetics:
 A:Gene: lysC
 C:Superfamily: aspartate kinase; aspartate kinase homology

C:Keywords: phosphotransferase

Query Match 21.0%; Score 590; DB 2; Length 461;
 Best Local Similarity 34.5%; Pred. No. 8,1e-34;
 Matches 165; Conservative 92; Mismatches 173; Indels 48; Gaps 15;

```

QY 84 SVYMKFGSSVSSAARMAEVAAGLITFPEEPPVVLSAMGKTNNLLA-----GKAV 137
DB 16 TVAKFGGTVAASFDASRSADVLSNPDLVILSASGIT--NLVLALADSGEPEKA 73
QY 138 GCGVHVSEIE--EMVMKSLHKTVDGLPXICNLSLVEQLKGIAMKE-----L 190
DB 74 D---HEDIRHIQYDIARTDSTVIR-----EIDRMENIMLSEASLST 118
QY 191 TPRSDYVVFGEQWSTRIFSAYLNKIRVAKROYDADIGFITTTDEFQNDILEATVPAY 250
DB 119 SPALDELVSNGELMSTLLFVELLRQAVAVEMVDVAKV-MRTNDRGRAEPDTSALAE 177
QY 251 AKRLHDMIQDPAIPVVTGFLGKMKSGAVTTLGRGSDLTATTIGALGREIQWKV 310
DB 178 AGTLAPRIED-ATVVTQGFISGK--GRITTLGRGSDYVTAALGELAVSRIDWTDV 235
QY 311 DGVLCDDNIYFHAKEVYLTFEEATELAYGAQVLPQSRPARBEDIPIRVKNSYNK 370
DB 236 PGITTDPRVYPAKRIIDKIAFEBAEMATGAKILHPATLAVNSDIFMGSSKDP 295
QY 371 AEGTLITQRMDXGLVLTSLVKSNTMDIVSTRMLQGYGLARVSGICYIEDLCIS 430
DB 296 AGTLVQNE---TYNPLFRALARRKQTLTLHLSLMLHARGLAEVFN--LASHSIS 350
QY 431 VDCVATSEVSVSIDSPKINSR--ELIOQASELDHVEELKEIAIYRLAQPAISLIG 488
DB 351 VDLITSEVSVALTIDITGSTSTGDSLTITS-----LITLSSICREVEDLALVAIG 405
QY 489 -NVQSSILIEKTRGVLRKSGVNTQMSQASKYNMILIYHSDPAALVEALHQAFFE 545
DB 406 NNLQACGVAKVEFVGL--DFNIRMICYGASHNLCELPBGNDADKRVQTLHYNLF 461
  
```

RESULT 6
 K12CD3
 aspartate kinase (EC 2.7.2.4) III, lysine-sensitive - Escherichia coli (strain K-12)
 N:Alternate names: aspartokinase III
 C:Species: Escherichia coli
 C>Date: 31-Mar-1988 #sequence_revision 17-Oct-1997 #text_change 25-Aug-2003
 C:Accession: G65209; A25659; F41098
 R:Bateman, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A.; Rose, D.J.; Mau, B.; Shao, Y. Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: G65209
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-449 <BLAT>
 A:Cross-references: GB:AE00475; GB:U00096; NID:g1790448; PIDN:AACT6994.1; PID:g17904
 A:Experimental source: strain K-12, substrain MG1655
 R:Casas, M.; Parsot, C.; Cohen, G.N.; Patte, J.C. J. Biol. Chem. 261, 1052-1057, 1986
 A:Title: Nucleotide sequence of lysC gene encoding the lysine-sensitive aspartokinase
 A:Reference number: A25659; MUID:86111734; PMID:3003049
 A:Accession: A25659
 A:Molecule type: DNA
 A:Residues: 1-57; C, 59-400; A, 402-449 <CAS>
 A:Cross-references: GB:M1812; NID:g146682; PIDN:AAA24095.1; PID:g146683
 A:Experimental source: K12
 R:Casas, M.; Ronceray, J.; Patte, J.C. Nucleic Acids Res. 11, 6157-6166, 1983
 A:Title: Nucleotide sequence of the promoter region of the E. coli lysC gene.
 A:Reference number: I41098; MUID:84015362; PMID:6312411
 A:Accession: I41098
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

QY 315 TCDPNITPAAKTVPYLTFEEATELAYFGAQLVHPQSMKPAEGDIPIRVKNSYNPKAPGT 374
 DB 228 TTDPRVVSAAKRIIDIAFEAAEWATFGAKVILHPATILPAVRSDIPIRVFVSGSKDPRAGGT 287
 QY 375 LITRQDMXGLVLTSLVLSKSNVTMLDIVSRMIGQVGFARVSGICYIEDLCISVDCV 434
 DB 288 LVCKNKE--NPPLEFRLALRNQTLTLHLSNMLHSGFLAEVFGILARN--ISTDLI 342
 QY 435 ATSEVSVSLDPSKIWSR--ELIQASELDHVEELEKAIIVLLQOPATISLGN-VE 491
 DB 343 TTSEVSVSLTDTGTSTGDTLLTQS-----LIMELALGRVEEBGLAVALLIGNDLS 397
 QY 492 QSLILEKTRVLRKSGVNVOMISOGASKNNMSLIVHSDPAKALVEALHQAFFE 545
 DB 398 KACGVKKEVFLVLEP--FNIRMICYGASHNLCFLVPGSDAEQVVKLHSLNLF 449

RESULT 9

AC1013
 aspartate kinase (EC 2.7.2.4) - Salmonella enterica subsp. enterica serovar Typh1 (strai
 C/Species: Salmonella enterica subsp. enterica serovar Typh1
 A/Note: this species has also been called Salmonella typh
 C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 25-Aug-2003
 C/Accession: AC1013
 R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A/Reference number: AB0502; MUID:21534947; PMID:11677608
 A/Accession: AC1013
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-449 <PAR>
 A/Cross-references: GB:AL51382; PIDN:CAD09204.1; PID:g16505208; GSPDB:GN00176
 C/Genetics:
 A/Gene: STY4416
 C/Superfamily: aspartate kinase; aspartate kinase homology
 C/Keywords: phosphotransferase

Query Match 20.1%; Score 566; DB 2; Length 449;
 Best Local Similarity 34.2%; Pred. No. 3.8e-32;
 Matches 162; Conservative 85; Mismatches 185; Indels 42; Gaps 14;
 QY 85 VNMKFGSSVSSAARAEVAGLITLTPERRPVVLSAMGKTTNNLLAGEKAVGCGVTHV 144
 DB 5 VAKFGGTIVADPDAMNRSDIYVSDANR-LVLSASAGITNLVALAE-----GMEPG 58
 QY 145 SEIEEMNVKSLHIKTVDLGLPXICNTSLYELBOLKGIAMKE-----LTPKSDLV 199
 DB 59 ERPAITDAIRKIOFDILRLRHP--NVIREIEIRLENTITLAEASLATSALTDLV 115
 QY 200 SFGGWSSTRIFSAVINKIRYKARQYADFITTFDEFGNAD-----ILEATYPAVAKRL 254
 DB 116 SHGELNSTLLFVELLRDVOAMFVPRKY-METSPRFGAEPDVAALAEAAQQLRL 174
 QY 255 HGMWIDDPALPVYTGFLGKMGKGAVTTLGRGSDLTATTIGKALGREIQWKCVDGVL 314
 DB 175 -----SETLVITQGFISSEK-GRITTLGRGSDYTAALLAEALMAAVDITWIVPGIY 227
 QY 315 TCDPNITPAAKTVPYLTFEEATELAYFGAQLVHPQSMKPAEGDIPIRVKNSYNPKAPGT 374
 DB 228 TTDPRVVSAAKRIIDIAFEAAEWATFGAKVILHPATILPAVRSDIPIRVFVSGSKDPRAGGT 287
 QY 375 LITRQDMXGLVLTSLVLSKSNVTMLDIVSRMIGQVGFARVSGICYIEDLCISVDCV 434
 DB 288 LVCKNKE--NPPLEFRLALRNQTLTLHLSNMLHSGFLAEVFGILARN--ISTDLI 342
 QY 435 ATSEVSVSLDPSKIWSR--ELIQASELDHVEELEKAIIVLLQOPATISLGN-VE 491
 DB 343 TTSEVSVSLTDTGTSTGDTLLTQS-----LIMELALGRVEEBGLAVALLIGNDLS 397

QY 492 QSLILEKTRVLRKSGVNVOMISOGASKNNMSLIVHSDPAKALVEALHQAFFE 545
 DB 398 KACGVKKEVFLVLEP--FNIRMICYGASHNLCFLVPGSDAEQVVKLHSLNLF 449

RESULT 10

F82328
 aspartate kinase III, lysine-sensitive VC0391 [imported] - Vibrio cholerae (strain N1696
 C/Species: Vibrio cholerae
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 25-Aug-2003
 C/Accession: F82328
 R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gmin, M.L.; Dodson, R.
 chardeon, D.; Esmolaeva, M.D.; Vamathavan, U.; Bae, S.; Qin, H.; Dragol, I.; Sellers
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.W.
 Nature 406, 477-483, 2000
 A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A/Reference number: AB2035; MUID:20406833; PMID:10952301
 A/Accession: F82328
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-479 <HEI>
 A/Cross-references: GB:AB004127; GB:AB003852; NID:9654808; PIDN:AAF93564.1; GSPDB:GN
 A/Experimental source: serogroup O1; strain N16961; biotype El Tor
 C/Genetics:
 A/Gene: VC0391
 A/Map position: 1
 C/Superfamily: aspartate kinase; aspartate kinase homology

Query Match 18.9%; Score 531; DB 2; Length 479;
 Best Local Similarity 32.6%; Pred. No. 1.2e-29;
 Matches 155; Conservative 95; Mismatches 183; Indels 42; Gaps 16;
 QY 86 VNMKFGSSVSSAARAEVAGLITLTPERRPVVLSAMGKTTNNLLAGEKAVGCGVTHV 145
 DB 34 VAKFGGTIVADPDAMNRSDIYVSDANR-LVLSASAGITNLVALAE-----GMEPG 58
 QY 146 EIEE-----NMVKSILHKTVDLGLPXICNTSLYELBOLKGIAMKE-----LTPKSD 196
 DB 85 STTRQALIQIADLHYIITDLPYVER---EVRKIVDSVSAEASFOSSKLTLD 141
 QY 197 YVSGECWSTRIFSAVINKIRYKARQYADFITTFDEFGNAD-----ILEATYPAVAKRL 254
 DB 142 HVAAGELISTYITLQWREBREGVDAFREDFEV-LRTDHDGRAEPQLKEITAOAKELV 200
 QY 255 HGMWIDDPALPVYTGFLGKMGKGAVTTLGRGSDLTATTIGKALGREIQWKCVDGVL 314
 DB 201 P--LCEQYVVVVGFTGSD-AQNTTTLGRGSDYSALAEAVQASGLEITWIVPGIY 256
 QY 315 TCDPNITPAAKTVPYLTFEEATELAYFGAQLVHPQSMKPAEGDIPIRVKNSYNPKAPGT 374
 DB 257 STDPRIAPKAAPIPEISFEBAEYANFGAKILHPETILPAVRHGIPVVGSSKEPEKGT 316
 QY 375 LITRQDMXGLVLTSLVLSKSNVTMLDIVSRMIGQVGFARVSGICYIEDLCISVDCV 434
 DB 317 WIRKQVE---SAPVRLALRNQTVTLRSQMHAGFLAKFEL--LAKHITSVDLI 371
 QY 435 ATSEVSVSLDPSKIWSRELIIQASELDHVE-ELEKAIIVLLQOPATISLGN-VEQ 492
 DB 372 TTSEIVSLTLDKT-----DTGGAPELPEAARABEELCEVEYVNCILALIGNKKE 426
 QY 493 QSLILEKTRVLRKSGVNVOMISOGASKNNMSLIVHSDPAKALVEALHQAFFE 547
 DB 427 SRGVAKQVSVL--EDYNLRMICYGASHNLCFLHESVSKQAVQKHKELFEB 479

RESULT 11

T39822
 aspartate kinase - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
 C/Accession: T39822
 R/Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Pohl, T.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 11:07:15 ; Search time 19 Seconds
(without alignments)
1534.699 Million cell updates/sec

Title: US-09-890-813-6

Perfect score: 2813
Sequence: 1 MALPVSAAAPRLVPSIRP.....QAFEDVLSQVEANLTVG 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt 42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	662	23.5	AK_METUA	Q57991 methanococ
2	569	20.2	AK3_ECOI	P08660 escherichia
3	516.5	18.4	AK_SCHPO	O60163 schizosacch
4	491.5	17.5	AK_YEAST	P10869 saccharomy
5	477.5	17.0	AKH_ECOI	P00861 escherichia
6	475.5	16.9	AKH_HAETN	P44505 haemophilus
7	463.5	16.5	AKH1_SERMA	P27725 serratia ma
8	461.5	16.4	AKH2_MAIZE	P49080 zea mays (m
9	452.5	16.1	AKH1_MAIZE	P49079 zea mays (m
10	440	15.6	AKH_BUCBP	P08674 buchnera ap
11	418	14.9	AK_THERM	P97151 thermus the
12	416.5	14.8	AK_COREP	O81911 corynebacte
13	416.5	14.8	AKH_BUCBP	Q8K309 buchnera ap
14	415.5	14.8	AK_CORGL	P26512 corynebacte
15	415.5	14.8	AK_MYCSM	P41403 mycobacteri
16	412.5	14.7	AK_CORFL	P41398 corynebacte
17	412.5	14.7	AKH_DAUCA	P37142 daucus caro
18	409.5	14.6	AK_MYCTU	P97048 mycobacteri
19	408.5	14.5	AK3_BACSU	P94417 bacillus su
20	402	14.3	AK2_BACST	P59229 bacillus sp
21	396	14.1	AK2_BACST	P53553 bacillus st
22	394	14.0	AKH_BUCAI	P57290 buchnera ap
23	380	13.5	AK1_BACSU	Q04795 bacillus su
24	374	13.3	AK_PBEAS	O06977 pseudomonas
25	368.5	13.1	AK2_BACSU	P08495 bacillus su
26	364.5	13.0	AK_RICPR	O92617 rickettsia
27	347	12.3	AK2H_ECOI	P00562 escherichia
28	344	12.2	AK_CHLIR	O84267 chlamydia t
29	340.5	12.1	AK_AQUAR	O67221 aquifex aeo
30	335	11.9	AK_HELPY	O92127 helicobacte
31	335	11.9	AK_HELPY	O25827 helicobacte
32	334.5	11.9	AK_CHIRP	O9210 chlamydia p
33	301	10.7	AK_CHIMO	O9p332 chlamydia m

34	139	4.9	231	1	PYRH_THEMA	Q9X10 thermotoga
35	120.5	4.3	1280	1	MDRI_LEIRN	O06034 leishmania
36	113.5	4.0	753	1	YEHU_ECOI	P75764 escherichia
37	110.5	3.9	661	1	MCFA_BACSU	P39214 bacillus su
38	108.5	3.9	608	1	DXS_THEMA	Q9X291 thermotoga
39	108	3.8	378	1	GRK_HAEIN	P44507 haemophilus
40	105	3.7	227	1	MYTH_SULSO	Q9782 sulfolobus
41	105	3.7	462	1	MUTL_CLOTT	O05491 clostridium
42	105	3.7	867	1	DPOL_RICPR	O05949 rickettsia
43	104.5	3.7	381	1	GRX2_ECOI	P23524 escherichia
44	104.5	3.7	468	1	SG3_HOMAN	O8KX2 homo sapien
45	104.5	3.7	1853	1	CIPA_CLOTM	Q06851 clostridium

ALIGNMENTS

RESULT 1
AK_METUA STANDARD; PRT; 473 AA.
ID AK_METUA
AC O57991;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable aspartokinase (EC 2.7.2.4) (Aspartate kinase).
GN M00571.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou J., Fleischmann R.D.,
RA Sutton G.G., Blake O., Olsen G.J., Clayton R.A., Gocayne J.D.,
RA Overbeek R., Kirkness E.F., Tomb J.F., Adams M.D., Reich C.I.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Rummel J.L., Nguyen D.,
RA Ullrich T.R., Kelley T.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Colton H.P., Fraser C.M., Smith H.O., Wessie C.R., Venter J.C.,
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii";
RL Science 273:1058-1073(1996).
CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-aspartate.
CC -!- PATHWAY: First step in the common biosynthetic pathway leading from Asp to diaminopimelate and Lys, to Met, to Ile and to Thr.
CC -!- SIMILARITY: Belongs to the aspartokinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL: U67506; AAB96565.1; --
CC PIR: C64371; C64371.
CC TIGR: M00571;
CC InterPro: IPR001048; AA_kinase.
CC InterPro: IPR002912; ACT.
CC InterPro: IPR005260; Asp_kin_monofn.
CC InterPro: IPR001341; Aspartate_kinase.
CC Pfam: PF00696; aak_kinase; 1.
CC TIGRfam: TIGR00656; asp_kin_monofn; 1.
CC TIGRfam: TIGR00657; asp_kinases; 1.
CC PROSITE: PS00324; ASPARTOKINASE; 1.
CC Transferrase; Kinase; Threonine biosynthesis; Complete proteome.
SQ SEQUENCE 473 AA, 51392 MW, 81166E0F456B5F CR64;

Query Match 23.5%; Score 662; DB 1; Length 473;
 Best Local Similarity 32.3%; Pred. No. 8.3e-39;
 Matches 160; Conservative 110; Mismatches 167; Indels 58; Gaps 11;

QY 83 LSVNKKFGSSSVSSAARNAEVAAGLLTPPERP--VYVLSAMGKTNLLAGEKAVCGC 140
 DB 1 MTTVMKFGSTVSGSERLHVAKITVKAKEDDDVYVVSAMSEVTLAVHISQAL--D 58
 QY 141 VHVSEIEEM-NMVKSLIKTVDEL-----GLPXICNTSLYELBQLKGIAMKELT 191
 DB 59 VNDIAKGVDFIKFIEKHAKYKALEBAIKSEIKYKIIDRIEIELEKVLIGVYLGELT 118
 QY 192 PRTSDYLVFEGECMSRIRSAVLNKRKARYDAFDIGFTTDEFGNA-----DILEAT 246
 DB 119 PMSRDYILFSGERLSSPIISGAIKRDGKSLMEGEGAGITIDNNFSAKYLEVERL 178
 QY 247 YPAVAKRLHGDWITDPAIPVVTGFLGKWKSGAVTTLGRGSDLTATTIGKALGLREIQV 306
 DB 179 LPL-----LKEGIIIPVTGFIGT--TEEGYITTLGRGSDYSAAIIGYGLADIIIEI 228
 QY 307 WQDVGVLTCDPNITPHAKTVYLTFEETELAYGAQVLPQSMRPARBGDIDVRYKNS 366
 DB 229 WIDVSGVYTTDPLVETARIRIKLSTIEAMELAYGAKVLPHTTEPAMEKGIITLVKNT 288
 QY 367 YNPKAPGLITRQDMDXGLVYLTSLKSNVTMLDITSTMIGQYGLARVSGICVIED 426
 DB 289 FESESGTILINDMMSDSIVAKISTI--KNVALINIGAGVVGSTPAARIFKALGEE 346
 QY 427 LCISVDCVATSEVSUSLDPKINSRELIQASLDRVLELEK----- 471
 DB 347 VNVILISQSSSETNIS-----LVSEEDVDKALKKKEFGDFPKKSFLLNNLT 394
 QY 472 IATVRLQORAITSLIG-NVEQSILLETGKGVLRKSGVNMISQASKYNMELIYHDS 530
 DB 395 IRVSVYDKVQCVISVAGAKMRGAKIGKIFITAVESGANNIMIQSSSEVNIISVIDEK 454
 QY 531 DAKALVALHQAFFE 545
 DB 455 DLINCVKRLHEKFE 469

RESULT 2
 AK3_ECOLI STANDARD; PRT; 449 AA.
 ID AK3_ECOLI
 AC P08660;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lysine-insensitive aspartokinase III (EC 2.7.2.4) (Aspartate kinase III).
 GN LYSC OR APK OR B4024.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA MEDLINE=8611734; PubMed=3003049;
 RA Cassan M., Parrot C., Cohen G.N., Patte J.-C.,
 RT "Nucleotide sequence of lysC gene encoding the lysine-sensitive
 RT aspartokinase III of Escherichia coli K12. Evolutionary pathway
 RT leading to three isofunctional enzymes.";
 RU J. Biol. Chem. 261:1052-1057(1986).
 RV [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RA MEDLINE=94089392; PubMed=8265357;
 RA Blatner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
 RA Daniels D.L.;
 RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
 RT region from 89.2 to 92.8 minutes.";

RL Nucleic Acids Res. 21:5408-5417(1993).
 RN [3]
 RP SEQUENCE OF 1-20 FROM N.A.
 RA MEDLINE=84015162; PubMed=6312411;
 RX Cassan M., Ronceteray J., Patte J.-C.;
 RT "Nucleotide sequence of the promoter region of the E. coli lysC
 RT gene.";
 RU Nucleic Acids Res. 11:6157-6166(1983).
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate
 CC -1- ENZYME REGULATION: SYNTHESIS AND ACTIVITY ARE SENSITIVE TO LYSINE,
 CC WHICH IS ONE OF THE END METABOLITES OF THE ASPARTIC ACID FAMILY
 CC BRANCHED PATHWAY.
 CC -1- PATHWAY: Lysine biosynthesis; first step.
 CC -1- SUBUNIT: Homodimer.
 CC -1- MISCELLANEOUS: ASPARTOKINASES I AND II ALSO CATALYZE THE SAME
 CC REACTION(S).
 CC -1- SIMILARITY: belongs to the aspartokinase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; M1812; AAA24095.1; -
 CC EMBL; U00006; AAC43118.1; -
 CC EMBL; AE000475; AAC76994.1; -
 CC EMBL; X00008; CAA24910.1; ALT_SEQ.
 CC PIR; G65209; KIECD3.
 CC Ecogene; Egi0550; lysC.
 CC InterPro; IPR001048; Aa_kinase.
 CC InterPro; IPR002912; ACT.
 CC InterPro; IPR005260; Asp_kin monofn.
 CC InterPro; IPR001341; Aspartate_kinase.
 CC Pfam; PF01842; ACT; 1.
 CC Pfam; PF01842; ACT; 1.
 CC Pfam; PF01842; ACT; 1.
 CC TIGRPFAM; TIGR00656; asp_xin_monofn; 1.
 CC TIGRPFAM; TIGR00657; asp_kinases; 1.
 CC PROSITE; PS00324; ASPARTOKINASE; 1.
 CC Transferrase; Kinase; Lysine biosynthesis; Complete proteome.
 CC DOMAIN 1 245 ASPARTOKINASE III.
 CC FT DOMAIN 246 449 INTERFACE.
 CC FT CONFLICT 58 58 G->C (IN REF. 1).
 CC FT CONFLICT 401 401 G->A (IN REF. 1).
 CC FT CONFLICT 401 401 G->A (IN REF. 1).
 CC SQ SEQUENCE 449 AA; 48531 MW; 5B41CB3A6E4D9848 CRC64;

Query Match 20.2%; Score 569; DB 1; Length 449;
 Best Local Similarity 34.0%; Pred. No. 2.4e-32;
 Matches 161; Conservative 89; Mismatches 182; Indels 42; Gaps 14;

QY 85 VNMKFGSSSVSSAARNAEVAAGLLTPPERP--VYVLSAMGKTNLLAGEKAVCGCIHV 144
 DB 5 VVSKFGSTVADPDMNRSADIVSDANVR-LVVLASAGITNLVLALE-----GLPEG 58
 QY 145 SEIEEMNMVKSILIKTVDELGLPXICNTSLYELBQLKGIAMKELT-----LTPRTSDYLV 199
 DB 59 ERREKIDAIRNTOFALIERLRYP---NVIREIERLLENITVLAAALATSPALTDDELV 115
 QY 200 SPEECMSTIRISAVLNKRKARYQVDADIGFTTDEFGND-----ILENTYAVAKRL 254
 DB 116 SHEELMSTLLFVEIIRREDVQAQWDPVAKV-WRTNDRGRREPDIALLAETALALQILPRL 174
 QY 255 HGDWIDPAIPVVTGFLGKWKSGAVTTLGRGSDLTATTIGKALGLREIQWQDVGV 314
 DB 175 -----NEGLVITQGFISENK-GRITTLGRGSDYTAALAEALHSRVDIWDVGIY 227
 QY 315 TCDPNITPHAKTVYLTPEATELAYFGAQLVLPQSMRPARBGDIPVRYKNSYNPKAPGT 374
 DB 228 TIDPRTVSAKRIIDEIAFAEAEMKATFAKVLHPATILLPAVSDIPFVVGSSKDPKAGGT 287

QY 375 LITRQDMDXGLVLTSLVSKSNMTMDIVSTRMIGQYGLARVSGICVIEDLCISVDCV 434
 DB 288 LVCNTE--NPPLPALAKRNQTLTLTSLNMLHSGFLAEVFGILARN--ISVDLI 342
 QY 435 ATSEVSVSLDPSKIMSR--ELIQASELDHVEELEKIAIVRLQORALISLIGN-VE 491
 DB 343 TISEVSVALTLDITSTSTGDTLLQS-----LMLELSALCRVEVEEGALVALIGNDLS 397
 QY 492 QSSLIEKTKRYLRSGVNVQVMSIGASAKVMSLIYHSDPAKALVELHQAFEE 545
 DB 398 KACGVKEVFGVLER--FNIRMICYAGSSHNLCFLVPSDEBAQVQKLSNLF 449

RESULT 3
 AK_SCHPO STANDARD; PRT: 519 AA.
 ID AK_SCHPO
 AC 060163;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable aspartokinase (EC 2.7.2.4) (Aspartate kinase).
 GN SPBC19FS.04
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_Taxid=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RC MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jasele K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Shelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymporter B.,
 RA Weidens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer C., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Roeder M., Galliard C., Tallada V.A., Garçon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez W., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Sipakovski G.V., Usery D., Barrell B.G., Nurse P.,
 RA "the genome sequence of Schizosaccharomyces pombe".
 RA Nature 415:871-880(2002).
 RL Nature 415:871-880(2002).
 RT
 CC -1- CATALYTIC ACTIVITY: ADP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- PATHWAY: First step in the common biosynthetic pathway leading
 CC from Asp to diaminopimelate, to Met, to Ile and to Thr.
 CC -1- SIMILARITY: Belongs to the aspartokinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL databank on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL022599; CAA18652.1; -
 CC DR F1R; T39822; T39822.

DR GenedB, Spombe, SPBC19FS.04; -
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR Pfam; PF00696; aakinese; 1.
 DR Pfam; PF01842; ACT; 2.
 DR TIGRfam; TIGR00657; asp_kinases; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 KW transferase; kinase; threonine biosynthesis.
 SC SEQUENCE 519 AA; 56812 MW; 0B392EB4AB8BCFB CRC64;
 Query Match 18.4%; Score 516.5; DB 1; Length 519;
 Best Local Similarity 28.9%; Pred. No. 1,4e-28;
 Matches 152; Conservative 107; Mismatches 186; Indels 81; Gaps 20;
 QY 85 VYMKFGSSVSS-AARMA-EVAGLILTFPEERFVVLAMGK-----TNNLLIGEXA 136
 DB 16 VYKFGSTVGKFP-KIADVAKYLS--TKRVALLVCSASTDTRKAGITRILIRTEAA 73
 QY 137 VCCGVINSEIENMMVMSLHKI---VDELG---LPXICNTSLYEQLKGIAMK 188
 DB 74 LRPVAGVHDLV--RIETDHVQARDFIQDVGIDELIDAFHADCVELDQYLAIRVLS 131
 QY 189 ELTPRTSDYLVSGECMSSTRIFSAYLNKIRYKARQYDAFDIGFTTD--EFGNADI--- 242
 DB 132 EVSPRTDLVIGMGBRSLSCRFMAVLKDQIGDS---EFIDMSHIDQREWRMLDASFYA 188
 QY 243 ----LEATYPAVAKRLHGMDIODBAIPVYTGFGKMGKGAVTTLGRGSDLTATTIGK 297
 DB 189 YLASGLASKVTVAGNK-----VVVVGFGFM--VPGGLSQIGRGTDPQCALILAV 237
 QY 298 ALGLREIOWKVDGVLCDFENIYPHAKTVYLTFFEBATELAYGAQVLHPOSRRPAREG 357
 DB 238 GLNABELQWKEVDGFIADPRKPTPALLLIFPEBAELTYGSEVHPFTMSGVYHA 297
 QY 358 DIPVRKVSYNKAPGTLI---TROR-----DMXGLV--VLISVLSKN 397
 DB 298 RIPRIKVNGNDRKGTIVFPDITSRHSGATPPHPKIMPDIDIASLANKGATVITKDT 357
 QY 398 VYMLDIVSRMLGQVGLFARVSGICVIEDLCISVDCVATSEVSVSLDPSKIMSRBLIQ 457
 DB 358 IMVINQSNRKLSAHGFLASIPAL--LKYKLAVDLITTSVHSMALYER-----S 407
 QY 458 QASELDHVEELEKIAIVRLQORALISLIGNVEGSSILIEKTRGV---LRKSGVNVQMI 514
 DB 408 DQGNHGAHVLELRRLGITDILHGLATLSLVGKHEMT--TSYAGMFCLEAEOINTEMI 465
 QY 515 SQGASAKVMSLIYHSDPAKALVELHQAFEDVLSQVEAE-NILIV 559
 DB 466 SQGASEINISCVIDEKMAVKALNVIHKELEPLALHEVPSQASMLV 511

RESULT 4
 AK_YEAST STANDARD; PRT: 527 AA.
 ID AK_YEAST
 AC P10869;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Aspartokinase (EC 2.7.2.4) (Aspartate kinase).
 GN HCM3 OR YER052C
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_Taxid=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=88115350; PubMed=2892836;
 RA Rafalski J.A., Falco S.C.;
 RA "Structure of the yeast HCM3 gene which encodes aspartokinase".
 RA J. Biol. Chem. 263:2146-2151(1988).
 RL [2]
 RN
 RP REVISIONS.

RA MEDLINE=90368723; PubMed=2168408;
 RA Rafalski J.A., Falco S.C.;
 RT "Structure of the yeast HOMO gene which encodes aspartokinase.",
 RL J. Biol. Chem. 265:15346-15346(1990).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=62886 / AB972;
 RA MEDLINE=9713264; PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Henessey K.M., Yelton M.A., Allen E.,
 RA Araujo R., Aviles E., Benno A., Brennan T., Carpenter J., Chen E.,
 RA Chery J.M., Chung E., Duncan M., Guman E., Hartzell G.,
 RA Hummel-Smith S., Hyman R.W., Kayser A., Kemp C., Laabhart D., Lew H.,
 RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oetner P.,
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome V,"
 RL Nature 387:78-81(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- PATHWAY: First step in the common biosynthetic pathway leading
 CC from Asp to diaminopimelate, to Met, to Ile and to Thr.
 CC -1- SIMILARITY: Belongs to the aspartokinase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL: J03526; AAA34681.1; -
 DR EMBL: U18796; AAB64587.1; -
 DR GERM: A35888; KTBXD.
 DR GERMONLINE: 139132; -
 DR SGD: S0000854; HOMO3.
 DR InterPro: IPR001048; Aa_kinase.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aak_kinase; 1.
 DR Pfam: PF01842; ACT; 2.
 DR TIGR: TIGR00657; asp_kinase; 1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 DR TRANSFERASE: Kinase; Threonine biosynthesis.
 KW SQUENCE 527 AA; 58109 MW; D4D28F8D4374898 CRC64;

Query Match 17.5%; Score 491.5; DB 1; Length 527;
 Best Local Similarity 28.1%; Pred. No. 7.7e-27;
 Matches 143; Conservative 107; Mismatches 180; Indels 79; Gaps 14;

QY 85 VNMKFGSSVSSAARAEVAGLILF-----PEERPVYVLSAMGK-----TTNNL--- 130
 DB 15 VVQKFGGTSGVKKFP-VQIVDDIVKHSKPDGPNNVAVCSARSSYTRAEGTSTLLKCC 73
 QY 131 -LAGEKAVGCGV-----HVSEIEEMNVKSLHTKTVDELGLFICNTSYEEFOLKKG 183
 DB 74 DLASQSESEFODILEVRQDHIDNADFLINPLQKLVND-----TKKELELYKXINA 127
 QY 184 IANMKELTPRTSDYLVSPGECNSTRIFSAVINKIRYKAROYDAFIFGFTTDEFGNADIL 243
 DB 128 SKVLGEVSRITVDMVASCBEKLSCLPMFALCDROGKAKY--VILSHIVSDPFSALD 184
 QY 244 EATYPAVAKELHGDW-----IODPAIVYTGFLGKMGKSGAVTTGSGSDPLTATITKA 298
 DB 185 NSRTFTIVQALKKRLAPFVSAKRIVPVGTGFFGL-VPTGLINGRGRTDUCALINVA 243
 QY 299 LGRREIQWMDVDVGLTCDPNIVPAKTVPYLTFFEBATELAYFGAQLVHPQSRPAREGD 358
 DB 244 VNDELQWKEVDGIFLADPRKVPPEARLIDSVTPBASLTYGSEVLIHPFMEQVIRAK 303
 QY 359 IPRVAKSYNPKAPGLI-----TRQDMXGGLVVLISI 392
 DB 304 IPIRIKVNQNLNGITLIPDNVAKKGESTPPHPENLSSSFYERKKG-----ATAI 356

QY 333 VLKSNVTMDIVSTRMLGQYGFLLARVSGICVIEDLCISYDCVATSEVSYSVSLDPKIMS 452
 DB 357 TTKNDIFVINIHNSKKTLKSHGFLAQIFTL--LDKXYMLVDLITSEVHVMALPIPDADS 414
 QY 453 RELIQASSELHVVSELEKIAIVLLOQRAIISLIG-NVBSQSLILEKTRGVLRKSGVNV 511
 DB 415 LKSLRQAE-----EKRIILGASVDITKKSLIVGVGMKQYIGAGTMTTLAEGINI 468
 QY 512 QMISGASKVMSLIVHSDQAKALVEALH 540
 DB 469 EMISGANEINISCVINESDSIKALQCIH 497

RESULT 5
 AKLH_ECOLI STANDARD; FRT; 820 AA.
 ID AKLH_ECOLI
 AC P00561; Q47659;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Bifunctional aspartokinase/homoserine dehydrogenase I (AKI-HDI)
 DE [includes: Aspartokinase I (EC 2.7.2.4); Homoserine dehydrogenase I
 DE (EC 1.1.1.3)].
 DE THRA OR THRA1 OR THRA2 OR B0002.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCBI_TaxID=562;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81077247; PubMed=7003595;
 RA Katinka M., Cossart P., Sibilli L., Saint-Girons I., Chalvignac M.A.,
 RA le Bras G., Cohen G.N., Yaniv M.;
 RT "Nucleotide sequence of the *thra* gene of *Escherichia coli*,"
 RL Proc. Natl. Acad. Sci. U.S.A. 77:5730-5733(1980).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizobuchi K., Nakata A.;
 RT "Systematic sequencing of the *Escherichia coli* genome: analysis of
 RT the 0-2.4 min region,"
 RL Nucleic Acids Res. 20:3505-3508(1992).
 RN (3)
 RP SEQUENCE FROM N.A.
 RX STRAIN=X12 / MG1655;
 RX MEDLINE=9534362; PubMed=7610040;
 RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
 RA Blattner F.R.;
 RT "Analysis of the *Escherichia coli* genome VI: DNA sequence of the
 RT region from 92.8 through 100 minutes,"
 RL Nucleic Acids Res. 23:2105-2119(1995).
 RN (4)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80077291; PubMed=390305;
 RA Cossart P., Katinka M., Yaniv M.;
 RT "Construction and expression of a hybrid plasmid containing the
 RT *Escherichia coli* *thra* and *thrb* genes,"
 RL Mol. Gen. Genet. 175:39-44(1979).
 RN (5)
 RP SEQUENCE OF 51-129.
 RX MEDLINE=80043179; PubMed=387092;
 RA Sibilli L., le Bras G., Cossart P., Chalvignac M.A., le Bras G.,
 RA Briley P.A., Cohen G.N.;
 RT "The primary structure of *Escherichia coli* K 12 aspartokinase I -
 RT homoserine dehydrogenase I: sequence of cyanogen bromide peptide CB
 RT 3,"
 RL Biochimie 61:733-739(1979).
 RN [6]
 RP REVISION TO 11.
 RX MEDLINE=83135751; PubMed=6298218;

RA Cossart P.;
 RA Unpublished results, cited by:
 RL Zakin M.M., Duchange N., Ferrara G.N.;
 RL J. Biol. Chem. 258:3028-3031(1983).
 CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P) (+) = L-aspartate 4-
 CC semialdehyde + NAD(P)H.
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- ENZYME REGULATION: The enzyme activities are regulated
 CC allosterically by L-threonine.
 CC -1- PATHWAY: Catalyzes 2 nonconsecutive reactions in the common
 CC biosynthetic pathway leading from Asp to diaminopimelate and Lys,
 CC to Met, and to Thr and Ile.
 CC -1- SUBUNIT: Homotrimer.
 CC -1- MISCELLANEOUS: Aspartokinase II-homoserine dehydrogenase II and
 CC aspartokinase II also catalyze the same reaction(s).
 CC -1- SIMILARITY: In the N-terminal section, belongs to the
 CC aspartokinase family.
 CC -1- SIMILARITY: In the C-terminal section, belongs to the
 CC homoserine dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V00361; CAA23660.1; ALT SRQ.
 DR EMBL; X68872; CAA48734.1; -
 DR EMBL; D10483; BAB96579.1; -
 DR EMBL; U14003; AAA97301.1; -
 DR EMBL; AE000111; AAC73113.1; -
 DR EMBL; V00360; CAA23659.1; -
 DR EMBL; M28570; AAA24673.1; -
 DR EMBL; M10644; AAA24671.1; -
 DR PIR; B64720; DEECK.
 DR EcoGene; EG10998; thra.
 DR InterPro; IPR001048; As_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR InterPro; IPR001342; Homoserine_dh.
 DR InterPro; IPR005106; NAD_Binding_3.
 DR Pfam; PF00696; aak_kinase; 1.
 DR Pfam; PF01842; ACT; 1.
 DR Pfam; PF00742; Homoserine_dh; 1.
 DR Pfam; PF03447; NAD_Binding_3; 1.
 DR TIGRfams; TIGR00657; asp_kinases; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 DR PROSITE; PS01042; HOMOSER DHEGENASE; 1.
 DR Transferrase; Kinase; Oxidoreductase; Threonine biosynthesis; NADP;
 KW Allosteric enzyme; Multifunctional enzyme; Complete proteome.
 FT DOMAIN 1 249 ASPARTOKINASE I.
 FT 1 250 470 INTERPACE
 FT DOMAIN 250 470 HOMOSERINE DEHYDROGENASE I.
 FT 1 250 470 820
 FT NP_BIND 471 478 NADP (POTENTIAL)
 FT 1 471 478
 FT CONFLICT 113 113 Q -> E (IN REF. 5).
 FT 1 113 113
 FT CONFLICT 230 230 D -> N (IN REF. 1, 2 AND 4).
 FT 1 230 230
 FT CONFLICT 375 375 Q -> L (IN REF. 1, 2 AND 4).
 FT 1 375 375
 FT CONFLICT 393 393 T -> A (IN REF. 1, 2 AND 4).
 FT 1 393 393
 FT CONFLICT 406 406 M -> A (IN REF. 1, 2 AND 4).
 FT 1 406 406
 FT CONFLICT 553 553 D -> N (IN REF. 1 AND 2).
 FT 1 553 553
 FT CONFLICT 583 583 TY -> IT (IN REF. 4).
 FT 1 583 583
 FT CONFLICT 587 588 DY -> I (IN REF. 1, 2 AND 4).
 FT 1 587 588
 FT CONFLICT 607 607 T -> R (IN REF. 1, 2 AND 4).
 FT 1 607 607
 FT CONFLICT 658 658
 FT 1 658 658
 FT CONFLICT 820 AA; 89119 MW; 08F28B9ECCAB10BD CRC64;
 FT 1 820 820
 FT SEQUENCE

Query Match 17.0%; Score 477.5; DB 1; Length 820;
 Best Local Similarity 29.4%; Pred. No. 13e-25;
 Matches 142; Conservative 98; Mismatches 210; Indels 33; Gaps 13;
 86 VAKFGGSVSSAPAAVAGLILFFPEERPV-VVLSAMGKTTNNLLAGEKAVG--CGVI 142

DB 3 VAKFGGTSVAAERLRLVADLIESNARQGVATVLSAPAKITNHLVMIEKTIISGQDALP 62
 QY 143 HUSELEEMWVKSLHIKTYDEIGLP-----XICNTSYLEQLKLGAMKELLPRRSDY 197
 DB 63 NISDAER-IFAEILTGLAAAGPFPALQKTFVQGEFAQGVHLSLGGCCPSINAA 121
 QY 198 LVSPFGCSTRIEFAVINKIRVAKQYDAFDIG-----FITDEF--GNADILEATYPAVA 251
 DB 122 LIRGKMSIALMAGL-----EARGHVTVIDPEKLLAAGHYLESIVDAESTRIIA 176
 QY 252 KRLHGMIDPAIPVVTGFLGKWKSGAVTTTGRGSGSLTTTGKALGLEIOWKVDVD 311
 DB 177 SRIPAD-----HVMVIMAGFTA-GNEKGELIWLGRNGSPSYSAVLAACIRADCCIMDVD 230
 QY 312 GVITCPNPIYPAKVTVPYTFEEAELEIYFGQVHLPQSPMPAREGDIPEVKNYSYKPA 371
 DB 231 GVTCPDPRQVPDRLKLSMSYQEMELSTFGAKVHPEHTTPIAQPICLINTGNPQA 290
 QY 372 PGLITRQDMXGLVLTSTIVLKSNTVMTLDIVSTRMLQYGFARVSGICYEDLCISV 431
 DB 291 PGLIGASRDEDE--LPVKGISNLNMMAFVSVCBGMKGMVAMARVAFAMSRARISVVL 348
 QY 432 DCVATSEVSVSLDPSK-INSRELIQASLELDHYVEELEKIAIVRLQRAIISLIGN- 489
 DB 349 ITGSSSEYSISFCVPSDQVRAERAMQEEFYELKEGLLEPLAVERL---AITSVADG 405
 QY 490 VEGSSLLIEKTRGVLRKSGVNTQMTSQGASKYMSLIVHSDAKALVALHQAFFEDVL 549
 DB 406 MRLRGISAKFPALARAINIVAIAGSSSEISIVVNNDDATGAVTHQMLFNTQV 465
 QY 550 SQV 552
 DB 466 IEV 468

RESULT 6
 AKH_HABIN STANDARD; PRT; 815 AA.
 AC P44505;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Bifunctional aspartokinase/homoserine dehydrogenase (AK-HD)
 DE [includes: Aspartokinase (EC 2.7.2.4); Homoserine dehydrogenase
 DE (EC 1.1.1.3)].
 DE THRA OR H10089.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxId=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; Pubmed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitznugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fitchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RA Rd.";
 RL Science 269:496-512(1995).
 CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P) (+) = L-aspartate 4-
 CC semialdehyde + NAD(P)H.
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- PATHWAY: Catalyzes 2 nonconsecutive reactions in the common
 CC biosynthetic pathway leading from Asp to diaminopimelate and Lys,


```

FT DOMAIN 250 470 INTERPACE (BY SIMILARITY).
FT DOMAIN 471 819 HOMOSERINE DEHYDROGENASE 1 (BY
FT SIMILARITY).
FT NP BIND 471 478 NADP (POTENTIAL).
FT VARIANT 330 330 G -> D (IN STRAIN HNR21; LOSS OF FEEDBACK
FT VARIANT 352 352 INHIBITION).
FT VARIANT 352 352 S -> F (IN STRAIN TLR156; LOSS OF
FT FEEDBACK INHIBITION).
FT VARIANT 479 479 A -> T (IN STRAIN HNR59; THR-RESISTANT
FT HDI).
SQ SEQUENCE 819 AA; 88494 MW; 1F18552B036AE39 CRC64;

Query Match 16.5%; Score 463.5; DB 1; Length 819;
Best Local Similarity 27.5%; Pred. No. 1.3e-24;
Matches 137; Conservative 111; Mismatches 186; Indels 65; Gaps 17;

QY 86 VMKEGSSVSAARMVAGLITFPEERPV-VILSAGKTNNTNLLAGEKAV-GCGVI-142
DB 3 VLFKGTSTVANAERFLRVADIMESNARQGVATVLSAPKINHLVAMIDKTVAGQDILP 62
QY 143 HVSEIE-----ENMNVKSLHKTVDYDELGLPYICNTSLYELBQLKQIA 185
DB 63 NMSDARIFADLLSLAQLPGFEYDLKGV-----VDQ-----EFAQIKQYLHGVS 109
QY 186 MMKELEPRISDVIVSGECMSTRIFSAVINKIRVAKQYDARDIG-----FITDFE--GN 239
DB 110 LUGQCPDVSNAATTCGEKLSAIMEGV-----PAKGYPTVIVPEKTLAQHYLEST 164
QY 240 ADILEATYPAVAKRELHGWIDOPALPVVTFGLKQKSGAVTTLGRGSDLTATTIGAL 299
DB 165 VDIAESTLRLAAAIAPAD-----HYLMAGFTA-GNDKELVVLGRNSDVSAAVLAACL 218
QY 300 GLREIQVMQDVQGLTCDPRITYPHATVYLFREETELAYGAQVLRQSRPAREDI 359
DB 219 RADCCETWDVDVYTCDPRTYDARLLKMSQOEAMELSTYGAVLHRTTTPAQPQI 278
QY 360 PVRVKNSYNPKAPGTILTR-QRDMGXGLVLTSLVSKSVNTLMDIVSTRMLQGYGLARV 418
DB 279 PCLIKTNSNPQARPTLIGDSTADMPVGINTL--NNMAININSGGMKGMVMAARV 335
QY 419 SGCIYBDELCTSDCAVTSFVSVSLVDSKTM-SRELIQQASEL--DHVELEIKAI 474
DB 336 FAVWSRAGISVAVITQSSSEYSSISFCVQSEIGRAFALKEEFYIELKQGVLDPLD--- 391
QY 475 VRLLOQRAITSLIGN-VEQSSILTEKTVLRKSGVNVOMISQASKVMMSLIYVDSAK 533
DB 392 --VMERLAIISVGDGKRTLRGISARFPALAAANINIVAIQSSSENSIVVYVNDSAT 449
QY 534 ALVEALHQAFEDDVLSQV 552
DB 450 TGVRVSHQMLENTDOVLEV 468

RESULT 8
AKH2 MAIZE STANDARD; PRT; 917 AA.
ID AKH2 MAIZE STANDARD; PRT; 917 AA.
AC P49080;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DB Bifunctional aspartokinase/homoserine dehydrogenase 2, chloroplast
DE Precursor (AK-HD 2) (AK-HSDH 2) [includes: Aspartokinase (EC 2.7.2.4);
DE Homoserine dehydrogenase (EC 1.1.1.3)].
GN AKHSDH2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OC NCBI_TaxID=4577;
RX [1]
RX SEQUENCE FROM N.A.
RX RP TISSUE=Seedling leaf;
RX MEDLINE=55148730; PubMed=7846152;

```

```

QY 220 KARQYDAPDIGFITTDEFGNADILEATYPAVAKELHGDWIODPA-IPVTFGLGKMGSG 278
DB 231 PCSWMDTREVLVVTSQGNQVD--PDYLECEKTLQKMFSPHAIIVANGFIAS--TAG 285
QY 279 AV-TTLGRGSGPLTTTIGKALGLREIQWVDVGLTCDPNIYFHAQVYLTFFEPATE 337
DB 286 NPTTLKRDGSDPSAIVGSLVPAQVITWTDVGFSAADPRKSEAVILTSLSYQWAME 345
QY 338 LAIFGAQVLPQSKPAREGDIPIRVKNSYNPKAPGTLITRQ----RDMXGLVLTST 392
DB 346 MPTFGANVLPRTIIPVKMDNPIYIRMTFSLAPGTWICQPANENGDD--ACVKSF 402
QY 393 VLKSNVTMDIVSTKYLQGYFLAVSGICIEDLCISYDCV--ATSEVSVSLDPSKI 450
DB 403 ATVDNIALVAVGEGTAGVPGTASAI--PSAVKDVGANVIMISQASSEHSVCFAV-PEK- 458
QY 451 MSRELIQOASEL-DHVEELE--KIAIVRLLOQRAIISLIG-----NVEQSLILEKTGR 502
DB 459 ---EVAVVAEHLDFPRELALAGRLSKVEVINGCSIIAAVGLRVASTPGVSAIIFD---- 511
QY 503 VLKSGVNTQMTSQASKVNSGLIVHSDAKALVEALHQAF 544
DB 512 ALAKANINRAIAGGSEYNTIVLKQDCVRAIIRAHSRPF 553

RESULT 9
AKH1_MAIZE
ID AKH1_MAIZE STANDARD; PRT; 920 AA.
AC P49079;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bifunctional aspartokinase/homoserine dehydrogenase 1, chloroplast
DE precursor (AK-HD 1) (AK-HSDH 1) [includes: Aspartokinase (EC 2.7.2.4);
DE Homoserine dehydrogenase (EC 1.1.1.3)].
GN AKHSDH1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACMAD clade; Panicoideae; Andropogonaceae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seedling leaf;
RX MEDLINE=95148730; PubMed=7846152;
RA Wehbauer G.J., Somers D.A., Matthews B.F., Gengenbach B.G.;
RT "Molecular genetics of the maize (Zea mays L.) aspartate kinase-
RT homoserine dehydrogenase gene family.";
RL Plant Physiol. 106:1303-1312(1994).
CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P) (+) = L-aspartate 4-
CC semialdehyde + NAD(P)H.
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
CC aspartate.
CC -1- PATHWAY: Catalyzes 2 nonconsecutive reactions in the common
CC biosynthetic pathway leading from Asp to diaminopimelate and Lys,
CC to Met, and to Thr and Ile.
CC -1- SUBUNIT: Homo- or heterodimer (Potential).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: In the N-terminal section; belongs to the
CC aspartokinase family.
CC -1- SIMILARITY: In the C-terminal section; belongs to the homoserine
CC dehydrogenase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, L33912; AAA74360.1; -.
CC PIR, T02953; T02953.

```

```

DR MaizeDB; 66609; -.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR001341; Aspartate_kinase.
DR InterPro; IPR001342; Homoserine_dh.
DR InterPro; IPR005106; NAD_binding_3.
DR Pfam; PF00696; aak_kinase; 1.
DR Pfam; PF01842; ACT; 2.
DR Pfam; PF00742; Homoserine_dh; 1.
DR Pfam; PF03447; NAD_binding_3; 1.
DR TIGRfam; TIGR00657; asp_kinases; 1.
DR PROSITE; PS00324; ASPARTOKINASE; 1.
DR PROSITE; PS01042; HOMOSER DHEGENASE; 1.
DR Transferrase; kinase; Oxidoreductase; Methionine biosynthesis; NADP;
KW Multifunctional enzyme; Chloroplast; Transic peptide;
KW Multigene family.
KW TRANSIT
KW CHAIN 1 92
FT 93 920
FT 93 341
FT 342 566
FT 567 920
FT 920 573
FT NP BIND 568 573
SQ SEQUENCE 920 AA; 100335 MW; 08DCF444BE45529 CRC64;

Query Match
Best Local Similarity 26.9%; Pred. No. 8,7e-24;
Matches 154; Conservative 115; Mismatches 239; Indels 65; Gaps 20;

QY 8 AARPRLVPSIPPPASSGHRGL-ACFGRTRGPGARGLSNVVADSTSRRAKQDGGV- 65
DB 13 AAFSTRPRLPHPPAAGDSTFRQCKWEKTCQSSFESSL-----RISRLPRTYH-GLILK 67
QY 66 -LGAVLGLGMEGL-----GDQLSVMKFGSGSVSSAARVAEYAGLITFPERPV 116
DB 68 NLARTPAGAVEQGEALADLPKDDWMS-VHKFGTCMGISERLHVADIVLADPBRKL 126
QY 117 VLSMGKTNNLLIAGEKAVGCVITVSELENNMKSLHITVDEL-----GLPXL 169
DB 127 VVVSAMSKYTTDMVTVLVKQAQRDSYIAVDE---VFDKMTYAKDLAGEDLARFLSQ 183
QY 170 CNTSLYELIOLKIGAMKELTPRTSDVLSFGCECMSTRIFSAVINKIRYKAOYADF 229
DB 184 LHADISNLKAMRAIYIGHATESPDSFVGHGLMAQLSVAIQKSGTCCSMWDTREV 243
QY 230 GFTTDEFGNADILEATYPAVAKELHGDWIODPALIPV-TGFLGKMGKSAVTTLGRGGS 288
DB 244 LVVNP---SGANQVDPDYLESEKLEKFSRCPAETIIATGFIH-STPENIPPTLKRDGS 299
QY 289 DLTATTTIGKALGLREIQWVDVGLTCDPNIYFHAQVLPYLTFFEBATELAFGAQVLP 348
DB 300 DPSAIIIGSLVAKRVYTTWTDVGFSAADPRKSEAVILTSLSYQWAMNSYGCANVLP 359
QY 349 QSMRPARBGDIPIRVKNSYNPKAPGTLITRQ----RDMXGLVLTSTVLKSNVTMDI 403
DB 360 RTIIPVKXNIPDIVRNIFNTSAPGTWICQPANENGLECYAFAPI---DIALVNV 416
QY 404 VETRMQLQGYFLAVSGICIEDLCISYDCV--ATSEVSVSLDPSKISRRELIQOASE 461
DB 417 EGTGAVAGVGTANALFGA--VADVAGANTIMISQASSEHSVCFAV-PEK---EVALVSA 469
QY 462 LDHVEE--LEKIAIVRLLOQRAIISLIGNVEQSLILEKTGRV-----LRKSGNV 511
DB 470 LHAEPREALAGRLSKVEVINGCSIIAIVG-----LRMASTPGVSAITLPDALAKANIN 523
QY 512 QMTSQASKVNSGLIVHSDAKALVEALHQAF 544
DB 524 RAIAQGCSEYNTIVLKQDCVRAIIRAHSRPF 556

RESULT 10
AKH1_BUCBP
ID AKH1_BUCBP STANDARD; PRT; 816 AA.

```

AC Q99AR4; 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Bifunctional aspartokinase/homoserine dehydrogenase (AK-HD)
 DE [includes: Aspartokinase (EC 2.7.2.4); Homoserine dehydrogenase (EC 1.1.1.3)]
 GN THRA OR BHP183;
 OS Buchnera aphidicola (subsp. *Baizongia pistaciae*).
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Buchnera.
 CC NCB1_TaxID=135842;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426901; PubMed=12522265;
 RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
 RA Bactolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
 RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
 RA "Reductive genome evolution in *Buchnera aphidicola*."
 RT Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
 CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P)(+) = L-aspartate + 4-semialdehyde + NAD(P)H.
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-aspartate.
 CC -1- PATHWAY: Catalyzes 2 nonconsecutive reactions in the common biosynthetic pathway leading from Asp to diaminopimelate and Lys, to Met, and to Thr and Ile.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- SIMILARITY: In the N-terminal section; belongs to the aspartokinase family.
 CC -1- SIMILARITY: In the C-terminal section; belongs to the homoserine dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE014016; AAC06915.1; -
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR InterPro; IPR001342; Homoserine_dh.
 DR InterPro; IPR005106; NAD_binding_3.
 DR Pfam; PF00696; aakinese; 1.
 DR Pfam; PF00742; Homoserine_dh; 1.
 DR Pfam; PF03447; NAD_binding_3; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 DR PROSITE; PS01042; HOMOSER_DHGEMASE; 1.
 DR TRANSIT; PS01042; HOMOSER_DHGEMASE; 1.
 KM Multifunctional enzyme; Complete proteome.
 FT DOMAIN 1 250 ASPARTOKINASE (BY SIMILARITY).
 FT DOMAIN 251 471 INTERFACE (BY SIMILARITY).
 FT DOMAIN 472 816 HOMOSERINE DEHYDROGENASE (BY SIMILARITY).
 FT NP_BIND 472 479 NADP (POTENTIAL).
 FT SEQUENCE 816 AA; 91918 MW; F8C3ADC0114576B5 CRC64;
 SQ
 Query Match 15.6%; Score 440; DB 1; Length 816;
 Best Local Similarity 26.8%; Pred. No. 5.5e-23;
 Matches 141; Conservative 96; Mismatches 182; Indels 108; Gaps 16;
 QY 86 VMKRGSSVSSAARAAPAGLI-LTFPERPVVLTSMGKTNNLLAGSKA-GGGVH 143
 DB 3 ILKRGSTLSNSELFFHVAITLNNNEQTAIVLSPGNTTLLSLAINTNNKNIIP 62
 QY 144 VSEIEEMNVKSLH-IKTVDLGLFXICN--TSLYELDQLKGIAMKELPRSDYI 198
 DB 63 IVQKIEKPKLKNINIVQEQLLEYKIKNNIEKMLKLVLLQGINVLKCPDKIRAKI 122
 QY 199 VSFECMSTRISAVLNKIRYKAROYDAFDIGFITT-----DEFGNADILERI 246

DB 123 ISSGEVLSISINSL-----ISGVNTTIIDPVKLLTKEDTYLNATV----- 166
 QY 247 YPAVAK-RLHGMIODPAIPVVTGLGKGMKSGAVTTIGRGSDLTATTIGKALGLEIQ 305
 DB 167 NIKISFRILSMKIPKPHHILMPGFTA-GNKGELVLTIGRNGSDYSATILSVCTNSTWCE 225
 QY 306 VMKDVGVLTCDPNIYPAKTPVLTFFESATELAYGAQVAPQPMRPARGDIPVRVKN 365
 DB 226 IMWDVGVTCDPKLVSDAKLTLSTSYNEALSTLGLKILHPNTIYIQKFKIPCTIKN 285
 QY 366 SYNPKAPGTLITQRDMXGLVLTSLVKSNTVMDIVSTRMLQGYFLARVSGICIE 425
 DB 286 TNPSSIGTKISCNHVNKNL-----ITGVYLE 314
 QY 426 DL-CIVDCVAVSEVSV-----SLDPSKIMRELIOAS----- 460
 DB 315 NVMPFISCLYSKNITTIIPKIFSCMSLSKIIITITQTSQNTISFCILKMTNTALHV 374
 QY 461 -----ELDHVELEKTAIVLLQGRALISLVNVSQSLILEKTRVLRKSGVNYQM 513
 DB 375 LHKALVDELKHKL-LKPIKVERKLLISVIS--SDIINNTKTEKVSILKHWVINTLA 430
 QY 514 ISQASGVMSLIVHSDPAKALVELHQAFFEDVDVLSQVEANLILVG 560
 DB 431 ISKGSKNSISIVKXDDILGVRLHKIEF-----KQCTAIFULIG 473
 RESULT 11
 AK_THETH STANDARD; PRT; 405 AA.
 ID AK_THETH
 AC P97151; F77991;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Aspartokinase (EC 2.7.2.4) (Aspartate kinase) [contains: Aspartokinase alpha subunit (ASK-alpha); Aspartokinase beta subunit (ASK-beta)].
 GN ASK OR ASKAB.
 OS Thermus thermophilus.
 CC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 CC Thermus.
 CC NCB1_TaxID=274;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX STRAIN=HB27;
 RX MEDLINE=99173891; PubMed=10074061;
 RX Kobasli N., Nishiyama M., Tanokura M.;
 RA "Aspartate kinase-independent lysine synthesis in an extremely thermophilic bacterium, *Thermus thermophilus*: lysine is synthesized via alpha-aminoadipic acid not via diaminopimelic acid."
 RT J. Bacteriol. 181:1713-1718(1999).
 RT [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 33923 / AT-62;
 RX MEDLINE=95291465; PubMed=773416;
 RX Nishiyama M., Kukimoto M., Beppu T., Horiuchi S.;
 RA "An operon encoding aspartokinase and purine phosphoribosyltransferase in *Thermus flavus*."
 RT Microbiology 141:1211-1219(1995).
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-aspartate.
 CC -1- PATHWAY: First step in the common biosynthetic pathway leading from Asp to diaminopimelate, to Met, to Ile and to Thr.
 CC -1- SUBUNIT: Tetramer consisting of 2 isoforms Alpha (catalytic) and 2 isoforms Beta (function not known) (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event: Alternative initiation:
 CC Comment=2 isoforms, Alpha/Aspartokinase alpha subunit (shown here) and Beta/Aspartokinase beta subunit, may be produced by alternative initiation;
 CC -1- SIMILARITY: Belongs to the aspartokinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB013311; BAA25848.1; -
 DR EMBL; AB013311; BAA25849.1; -
 DR EMBL; D37928; BAA07146.1; -
 DR EMBL; D37928; BAA07147.1; -
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR005260; Asp_kin monofn.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR Pfam; PF00696; aak_kinase; 1.
 DR Pfam; PF01842; ACT; 2.
 DR TIGRfams; TIGR00656; asp_kin monofn; 1.
 DR TIGRfams; TIGR00657; asp_kinase; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 KW Transferase; Kinase; Diaminopimelate biosynthesis;
 KW Lysine biosynthesis; Alternative initiation.
 FT CHAIN 1 405 ASPARTOKINASE, ISOFORM ALPHA.
 FT CHAIN 245 405 ASPARTOKINASE, ISOFORM BETA.
 FT INIT MET 245 245 FOR ISOFORM BETA.
 FT VARIANT 126 126 E -> D (IN STRAIN AT-62).
 SQ SEQUENCE 405 AA; 43333 MW; 06F7B79E9C8960EA CRC64;

Query Match 14.8%; Score 418; DB 1; Length 405;
 Best Local Similarity 27.9%; Pred. No. 7,3e-22;
 Matches 135; Conservative 88; Mismatches 149; Indels 112; Gaps 18;

QY 85 VVMKFGSSVSAAARMAVAGLILTPPE-REPVLVLSAMKNTNNLLAGEKAVGCVI 142
 4 VVQKGGTSVGDLEHKKVQRIAHYREKGRHLAVVSAHGHT----- 47
 DB 143 HVSEIEEMNWKSLHKTVDLGLPXCNTSLYELQLLKGIAMKELTPRTS----DYL 198
 48 -----DEL-----TALAKRVAPRPPELDLL 69
 DB 199 VSPGECWTRIFSAVLNKKIRKARQYDADPFGITTFDEFGNADILEATPAVAKRLHDM 258
 70 TTTEGQVSVALLMQMLWAGIPAKGVQVQIGITTDGRGDAHILEVNPARIREL---- 125
 QY 259 IQDDPAIVVTGFLGKWKSGAVTTLRGSSDLTATTIGALGLREIQWKKVDVGLTCDP 318
 126 -EQGFVAVIAGFMGT-TPEGEITTLRGSGDTTAAVAIAALAKCEIYTTDEGYTTDP 183
 QY 319 NIYPHAKTVPYLTFEEATELAFGAOVLPQSMRPARBGDIFVRVNSNYPKAPGLITR 378
 184 HILPEARKLSVIGDQMLEMAAGAVLHPRAVYAKRGVLAHRSSTSYN-PTLVYKE 242
 DB 379 -QSDMDXGLVLTLSVYKSNVTMLDIVSRMG--QYGFLLARVSGICYEDLCISVDCV 434
 243 VVAMEMDK--AVTGVALD-----LDHAQIGLIGIPQPIAAKV-----FQ 280
 QY 435 ATSEVSVSVL-----DPSK-----IMSRELIOQASE-LDHVEELEKIAIVRLIQ 480
 DB 281 ALMERGLAVMTIQGVGHPDSKQMAFYKCDPROEALALEPLAELGSGAIIIR--PD 338
 QY 481 RAISLIG-NVEQSSILEKTVLRKSGVNVQMTISQASKVNSLIVHSDAKALVAL 539
 DB 339 IAKYSIVGLASTPEVPAPMFQAVASTGANIEMIR--TSEVRISVIIIPAEVAAALRAV 366
 QY 540 HQAF 543
 DB 397 HQAF 400

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Aspartokinese (EC 2.7.2.4) (Aspartate kinase) [Contains: Aspartokinese
 DE alpha subunit; Aspartokinese beta subunit].
 GN LYSC OR ASK OR CEO220.
 OS Corynebacterium efficiens.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=152794;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Y8-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RA Itaya H., Kimura E., Kawahara Y., Sugimoto S.;
 RT "Ysc of Corynebacterium efficiens."
 RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Y8-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RX MEDLINE=2273752; PubMed=12840036;
 RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
 RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikee K.,
 RA Gotohori T.;
 RT "Comparative complete genome sequence analysis of the amino acid
 RT replacements responsible for the thermostability of Corynebacterium
 RT efficiens."
 RL Genome Res. 13:1572-1579(2003).
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- ENZYME REGULATION: Feedback inhibition by lysine and threonine.
 CC -1- PATHWAY: First step in the common biosynthetic pathway leading
 CC from Asp to the cell wall precursor meso-diaminopimelate, to Lys,
 CC to Met, to Ile and to Thr.
 CC -1- SUBUNIT: Tetramer consisting of 2 isoforms Alpha (catalytic) and 2
 CC isoforms Beta (function not known).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative initiation;
 CC Comment=2 Isoforms, Alpha/Aspartokinese alpha subunit (shown
 CC here) and Beta/Aspartokinese beta subunit, are produced by
 CC alternative initiation;
 CC -1- SIMILARITY: Belongs to the aspartokinese family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB083129; BAA88820.1; -
 DR EMBL; AP005214; BAC17030.1; AUT_INIT.
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR005260; Asp_kin monofn.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR Pfam; PF00696; aak_kinase; 1.
 DR Pfam; PF01842; ACT; 2.
 DR TIGRfams; TIGR00656; asp_kin monofn; 1.
 DR TIGRfams; TIGR00657; asp_kinase; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 KW Transferase; Kinase; Diaminopimelate biosynthesis;
 KW Lysine biosynthesis; Alternative initiation; Complete proteome.
 FT CHAIN 1 421 ASPARTOKINASE, ISOFORM ALPHA.
 FT CHAIN 250 421 ASPARTOKINASE, ISOFORM BETA.
 FT INIT MET 250 250 FOR ISOFORM BETA.
 SQ SEQUENCE 421 AA; 44793 MW; F0D63030693CB5F CRC64;

Query Match 14.8%; Score 416.5; DB 1; Length 421;
 Best Local Similarity 28.1%; Pred. No. 9.9e-22;
 Matches 134; Conservative 88; Mismatches 162; Indels 93; Gaps 14;

QY 85 VVMKFGSSVSAAARMAVAGLILTPPE-REPVLVLSAMKNTNNLLAGEKAVGCVI 142

```

Cc 143 HSEIEEMNNVKS.IHKTVDLGLPXICNTSLYEELQKGIAMKELTPTSDYLSFG 202
Cc 52 -----DLAAVNPVPAREM-----DWLITAG 73
Cc 203 ECKSTRIFSAVINKIRKARQYAPDIFGITTDGEMNDILATYPAKRLHGMWDQD 262
Cc 74 ERTSNALVAVALLSLGAEQSFSGQAVLTERGHNATIVVTGRRREAL-----DEG 128
Cc 263 AIPVYTGFLGKMGKSAVTTLGRGSDLTATTIGKALGRIQWKNVDYGVLTCDPNITP 322
Cc 129 KICIVAGFQGVNKNETDVTTLIGRGSDTLTAVLAALADGVCIEISVDGYVTADPRIVP 188
Cc 323 HACTVPLTTEETELAYGAQVLTQHSRPARREGIPRVKNSYNPKAPGLLI-RORD 381
Cc 189 NAKRLERLSEEWLELAAVSKTLVRSEVYAAAFVPRKRSYS-NDPGTLTSGSWED 247
Cc 382 MDXGLVLTSTIVL---KSNVTMDIVSTMLGQYGLFARVSGICYIEDLCISVDCV--- 434
Cc 248 IPWEAVLTGVALDKSEAKTVLIGPD-----KPGRAKV--FRALADAEIWDVQLNV 300
Cc 435 -----ATSEVSVV--SLDPSKWSRELLIQASELDHVELEKAIYRLLOQRAIILI 487
Cc 301 SSVEDGTTDITFTCPRSDGPRAMELKMKQOQGDWNTVLVD-----DQVGKSVLV 350
Cc 488 G-NVEQSLILEKTGRVLRKSGVNVQMGASKVMNSLIHSDPAKALVEALHQA 543
Cc 351 GAGKSHPGVTAFEMALRDVNNVLLIS--TSEIRISVLIREDDLDKSAKALHKEF 405

RESULT 13
AKH_BUCAP STANDARD; PRT; 814 AA.
ID AKH_BUCAP STANDARD; PRT; 814 AA.
AC Q8K9U;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bifunctional aspartokinase/homoserine dehydrogenase (AK-HD)
DE [includes: Aspartokinase (EC 2.7.2.4); Homoserine dehydrogenase (EC 1.1.1.3)].
CN THRA OR BUSC188.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
CX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tamas I., Klasson L., Canback B., Naeslund A.K., Eriksson A.-S.,
RA Werngreen U.J., Sandstrom J.P., Moran N.A., Andersson S.G.E.,
RA "50 million years of genomic stasis in endosymbiotic bacteria.";
RL Science 296:2376-2379 (2002).
CC -1 CATALYTIC ACTIVITY: L-homoserine + NAD(P)(+) = L-aspartate 4-
CC semialdehyde + NAD(P)H.
CC -1 CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
CC aspartate.
CC -1 PATHWAY: Catalyzes 2 nonconsecutive reactions in the common
CC biosynthetic pathway leading from Asp to diaminopimelate and Lys,
CC to Met, and to Thr and Ile.
CC -1 SUBUNIT: Homotetramer (By similarity).
CC -1 SIMILARITY: In the N-terminal section; belongs to the
CC aspartokinase family.
CC -1 SIMILARITY: In the C-terminal section; belongs to the homoserine
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/

```

```

Cc or send an email to license@isb-sib.ch).
Cc EMBL: AB014094; AAM67753.1; -.
Cc InterPro: IPR001048; AA_kinase.
Cc InterPro: IPR002912; ACT.
Cc InterPro: IPR001341; Aspartate_kinase.
Cc InterPro: IPR001342; Homoserine_dh.
Cc InterPro: IPR005106; NAD_binding_3.
Cc Pfam: PF00696; aak1nase; 1.
Cc Pfam: PF01842; ACT; 2.
Cc Pfam: PF00742; Homoserine_dh; 1.
Cc Pfam: PF03447; NAD_binding_3; 1.
Cc TIGRfam: TIGR00657; asp_kinases; 1.
Cc DR PROSITE: PS00324; ASPARTOKINASE; 1.
Cc DR PROSITE: PS01042; HOMOSER DHEGENASE; 1.
Cc KM Transferrase; Kinase; Oxidoreductase; Threonine biosynthesis; NADP;
Cc Multifunctional enzyme; Complete procenome.
Cc FT DOMAIN 1 248 ASPARTOKINASE (BY SIMILARITY).
Cc FT DOMAIN 249 469 INTERPACE (BY SIMILARITY).
Cc FT DOMAIN 470 469 HOMOSERINE DEHYDROGENASE (BY
Cc SIMILARITY).
Cc FT NP_BIND 470 477 NADP (POTENTIAL).
Cc SQ SEQUENCE 814 AA; 91378 MW; 162CDD52590D1A0C CRC64;

Query Match 14.8%; Score 416.5; DB 1; Length 814;
Best Local Similarity 25.7%; Pred. No. 2,4e-21;
Matches 127; Conservative 111; Mismatches 197; Indels 59; Gaps 14;

Cc 86 VMKFGSSVSAAAEVAGLI-LTPPEPRPVVYLSAMGKTNNLLAGEKAVGCVIHW 144
Cc 3 LKFGSTLANAKKFLCVADIEKKKKEQIAVVASAPAKITNLTATIEKKIDDEV--- 59
Cc 145 SEIEEMNNVKS.IHKTVDLGL---LXICNTSYELE-----QLKGIAMKELTPT 194
Cc 60 --LKKINLANKNFIELIDIKRIQPLPYENTKSTIELEPKKIINGILLILQCEGI 117
Cc 195 SDYLSFEGMSRTPSAVINKIRKARQYAPDIFGITTT-DEFGN-----ADILEATYP 248
Cc 118 KPITIRGELISVDIMKNTL-----QSNHETILNPNTNLISGNYLSDITDKES--- 169
Cc 249 AVAKRLHGMIDOPALPVVTGVLGQWMSGAVTTLGRGSDLTATTIGKALGRIEQW 308
Cc 170 --KRIKIKINIDQKNTILMAGFIA-GNKEGEVLVGRGSDYSAIILASCNAKCEITWT 226
Cc 309 DVDGVLTCDPNITPFAKTVPLTTEETELAYGAQVLTQHSRPARREGDIPRVKNSYN 368
Cc 227 DVDGVLTADPRIVSNLYLDIYSQEAHELSPFAKVLHPRITIEISQFQIPCVIKNTNN 286
Cc 369 FKAPGLITRQDMXGLVLTSTIVLSKSNVTMDIVSTRMLGQYGLFARVSGICYIED 428
Cc 287 TESKGTWIGKENNPSDN--SLKGVTYLDIIMFNISGCLDQSGNTIARIFTILSRE 344
Cc 429 ISVDCAVTSVSVSVSLDPSKWSRELLIQASELDHVELEK-----IAIVRL 478
Cc 345 IILIGSSSENQVN-----FCFEEKDIYIILILKETTLEIKSLGNDPRIV 392
Cc 479 QCARLISLTIG-NVEQSLILEKTGRVLRKSGVNVQMGASKVMNSLIHSDPAKALVE 537
Cc 393 KNLTILISTVGSINSEKNTASKIFSSLSKINVLALAHGSKHSISIVIKKENLLQGIQ 452
Cc 538 ALHQA-PPEDDVLS 550
Cc 453 NINHTLFFKXTIIN 466

RESULT 14
AK_CORGL STANDARD; PRT; 421 AA.
ID AK_CORGL STANDARD; PRT; 421 AA.
AC P26512; G59286;
DT 01-AUG-1992 (Rel. 23, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Aspartokinase (EC 2.7.2.4) (Aspartate kinase) [contains: Aspartokinase

```

DE alpha subunit; Aspartokinase beta subunit].
GN LYSC OR CGJ0251.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX MEDLINE=92065816; PubMed=1956296;
RA Kallnowski U., Cremer U., Bachmann B., Eggeling L., Sahn H.,
RA Puhler A.,
RT "Genetic and biochemical analysis of the aspartokinase from
RT Corynebacterium glutamicum";
RL Mol. Microbiol. 5:1197-1204(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.,
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032";
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-51 FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX MEDLINE=94161495; PubMed=8117072;
RA Patek M., Kumbach K., Eggeling L., Sahn H.,
RT "Leucine synthesis in Corynebacterium glutamicum: enzyme activities,
RT structure of leuA, and effect of leuA inactivation on lysine
RT synthesis";
RL Appl. Environ. Microbiol. 60:133-140(1994).
RN [4]
RP SEQUENCE OF 158-421 FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX MEDLINE=91094767; PubMed=1980002;
RA Kallnowski U., Bachmann B., Thierbach G., Puhler A.,
RT "Aspartokinase genes lysC alpha and lysC beta overlap and are
RT adjacent to the aspartate beta-semialdehyde dehydrogenase gene and in
RT Corynebacterium glutamicum";
RL Mol. Gen. Genet. 224:317-324(1990).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
CC aspartate.
CC -1- ENZYME REGULATION: Feedback inhibition by lysine and threonine.
CC -1- PATHWAY: First step in the common biosynthetic pathway leading
CC from Asp to the cell wall precursor meso-diaminopimelate, to Lys,
CC to Met, to Ile and to Thr.
CC -1- SUBUNIT: Tetramer consisting of two isoforms Alpha (catalytic) and
CC two isoforms Beta (function not known).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative initiation;
CC Comment=2 isoforms, Alpha/Aspartokinase alpha subunit (shown
CC here) and Beta/Aspartokinase beta subunit, are produced by
CC alternative initiation;
CC -1- SIMILARITY: Belongs to the aspartokinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X57226; CAA40502.1; -
CC EMBL: X57226; CAA40503.1; -
CC DR EMBL: AP005274; BAB97644.1; -
CC EMBL: X70959; CAA50296.1; ALT_SEQ.
CC PIR: I40723; I40723.
CC PIR: S15276; S15276.
CC InterPro: IPR001048; Aa_kinase.
CC InterPro: IPR002912; ACT.
CC InterPro: IPR005260; Asp_kin_monofn.
CC InterPro: IPR001341; Aspartate_kinase.
CC Pfam: PF00696; aakinae; 1.

DR Pfam: PF01842; ACT; 2.
DR TIGRFAMs: TIGR00656; asp_kin_monofn; 1.
DR TIGRFAMs: TIGR00657; asp_kinase; 1.
DR PROSITE: PS00324; ASPARTOKINASE; 1.
KW Transferase; Kinase; Diaminopimelate biosynthesis;
KW Lysine biosynthesis; Alternative initiation; Complete proteome.
FT CHAIN 1 421
FT INIT MET 250 421
FT MUTAGEN 301 301
FT CONFLICT 40 40
SQ SEQUENCE 421 AA; 44754 MW; E36BAD0061D50827 CRC64;
Query Match 14.8%; Score 415.5; DB 1; Length 421;
Best Local Similarity 29.0%; Pred. No. 1.2e-21;
Matches 137; Conservative 84; Mismatches 167; Indels 85; Gaps 14;
QY 85 VMKRGSSVSAAAEVAGLITPE--ERPVTLSAMGKTTNNLLAGEKAVGCVI 142
DB 4 VQKIGSSLSAARINVAERIVATKAQNDVVVCSMGDTTDL----- 51
QY 143 HVSEIEENNVKSLHIKTVDLGLPKICNTSLVELQLKGLAMKELTPRTSDYVSG 202
DB 52 -----ELAAAVPVPAREM-----DMLLTAG 73
QY 203 EGMSTRISAVINKIRKARQYDAFDIGFITTDFGADILEATYRAVKRLHGDWIODP 262
DB 74 ERISALVAKMAIESLGAEDSFTGSQAGVLTPEHGRARIVDTTPGRVREAL-----DEG 128
QY 263 AIPVVTGFLGKQWKSQVTTLGRGSDLTATITGKALGREIOWKDVGVLTCDNITYP 322
DB 129 KICIVAGQGVNKRPTDTTLGRGSDTLVALAALNADVCIEYSDVGVYADPRIVP 188
QY 323 HAKTVPYLTPEATELAFQAQVIVHPQSMAPAREGDIPIPVKSNYPKAPGLIT-RQD 381
DB 189 NAKIKETKISFEEMLEAVGSKILIVSVEARFNPFLKRSYS-NDKGILLASGMD 247
QY 382 MDXGLVLTSLIVL--KSNVTMLDIVSTRMLQGVFLARYSGICYIEDLCISVQVATSE 438
DB 248 IVEEAVLTGVALTKSEKVTVLGISD-----KGEAAKV--PALADAEINIDWLVQNV 300
QY 439 VSV-SVSDPSKISRELIOQASLDHVELEKIAL-----VRLIQGAILSLIG-IV 490
DB 301 SEVEDGTDITFTCPRSQGRAMEI-----LKKLQVQGWNTLVYDDQVKSLVQAGM 354
QY 491 EGSLLLEKTRVLRKSGVNVQMTISQASKVNLSLIYHSDAKALVEALHQA 543
DB 355 KHPGVTAEPFHEALRDVNVNIELIS--TSERISVLIRREDDLDAAARALHEGF 405
RESULT 15
AK MYCSM STANDARD; PRT; 421 AA.
ID AK MYCSM
AC P41403;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Aspartokinase (EC 2.7.2.4) (Aspartate kinase) [Contains: Aspartokinase
DE alpha subunit (ASK-alpha); Aspartokinase beta subunit (ASK-beta)].
GN ASK.
OS Mycobacterium smegmatis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 607 / mc(2)6 / NRRL B-692;
RX MEDLINE=94254720; PubMed=7910936;
RA Cirillo J.D., Weisbrod T.R., Pascopella L., Bloom B.R.,
RA Jacobs W.R. Jr.,
RT "Isolation and characterization of the aspartokinase and aspartate
RT semialdehyde dehydrogenase operon from mycobacteria";

RL Mol. Microbiol. 11:629-639(1994).
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- ENZYME REGULATION: Feedback inhibition by lysine and threonine (By
 CC similarity).
 CC -1- PATHWAY: First step in the common biosynthetic pathway leading
 CC from Asp to the cell wall precursor meso-diaminopimelate, to Lys,
 CC to Met, to Ile and to Thr.
 CC -1- SUBUNIT: Tetramer consisting of 2 isoforms Alpha (catalytic) and 2
 CC isoforms Beta (function not known) (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative initiation;
 CC Comment=2 isoforms: Alpha/Aspartokinase alpha subunit (shown
 CC here) and Beta/Aspartokinase beta subunit, may be produced by
 CC alternative initiation;
 CC -1- SIMILARITY: Belongs to the aspartokinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z17372; CAA78984.1; -.
 DR EMBL: Z17372; CAA78985.1; ALT_INIT.
 DR PIR: S42422; S42422.
 DR InterPro: IPR001046; Aa_kinase.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR005260; Asp_kin_monofn.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aak:kinase; 1.
 DR Pfam: PF01842; ACT; 2.
 DR TIGRfams: TIGR00636; asp_kin_monofn; 1.
 DR TIGRfams: TIGR00657; asp_kinases; 1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 KM Transferase; Kinase; Diaminopimelate biosynthesis;
 KM Lysine biosynthesis; Alternative initiation.
 FT CHAIN 1 421 ASPARTOKINASE, ISOFORM ALPHA.
 FT INT MET 250 421 ASPARTOKINASE, ISOFORM BETA.
 FT INT MET 250 250 FOR ISOFORM BETA.
 SQ SEQUENCE 421 AA; 44458 MW; 92689FC62520CD CRC64;

 Query Match 14.8%; Score 415.5; DB 1; Length 421;
 Best Local Similarity 27.5%; Pred. No. 1.2e-21;
 Matches 134; Conservative 87; Mismatches 152; Indels 115; Gaps 16;

 QY 85 VMKFGSSVSAAKMAEVAAGLITPE--ERPVTLSMGKTTNLLAGKAVGCGVI 142
 DB 4 VVQKYGSSVADAEIRIRVAERIVETKAGNDVVAVVSAAGDTDDL----- 51
 QY 143 HVSEIEMNWKSLHKTVDLGLPKICNTSLYEQLKGLAMKELTPRTSDYLVSPG 202
 DB 52 -----DLARQVSPAP--PPREMDMLTAG 73
 QY 203 EGMSTRIFSAVINKIRYAKOYDAFDIGTTDFGNAIDLENTYPAVAKRLHGMIDP 262
 DB 74 ERIENALVMAIEISLGAQARSFTGSAQAVITTTGHNAKIIDVT---PGLRL-DALDEG 128
 QY 263 AIPVVTGLGKWSGAVTTLGRGSDLTATTIGKALGLEIOWMKVDVGLTCDPNIYP 322
 DB 129 QIVVAGFQSVQSQSKVTTIGRGSDTTAVAAALDADVCEIYTDVGIPTADPRIVP 188
 QY 323 HAKTVPIYTFEEATELAYFQAQVLAHQSMRPABGDIPIRVKSNYNPKAPGTLITRQ-RD 381
 DB 189 NARHLDIVSFEEMLEMAACGAKYLMRCVEYARYNPIHVRSSYSYDK--PGITVKSIED 247
 QY 382 MDXGLVLTSTV--LKSNTMDIVSTRMLGQGFARVSGICYIEDLCISVDVATSE 438
 DB 248 IPMEDAILTGVAHRSSEAKTVVGLPDVP-----GYAAKYFR-----AVAE 288
 QY 439 VSVSVSLDPSKIMRELIQASELD-----HVEEL---EKIATVRLI 478

Db 289 ADVNIDM-----VQNISKIEDKTDITFCARDNGPRAVERKLSALKSEIGFQVL 339
 QY 479 QQRAT--ISLIGNEQSSLITEKT--GRVLRKSSGVNVOMISQASKNMSLIVHDSDAKL 535
 Db 340 YDDHIGKVLIGAGWRSHPGVTATFCALAEAGINIDLI--TSEIRISVLIKDTEDKA 397
 QY 536 VEALHOAF 543
 Db 398 VSALEHAP 405

Search completed: March 23, 2004, 13:16:13
 Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:00:27 ; Search time 71 Seconds
(without alignments)
2468.594 Million cell updates/sec

Title: US-09-890-813-6

Perfect score: 2813
Sequence: 1 MAIPVRSAAAPRLVPSTIP.....QAFEDVLSQVEANILV 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1881	66.9	510	10	Q8LPZ8
2	1880	66.8	565	10	Q851Z6
3	1868.5	66.4	581	10	Q8RYL8
4	1835	65.2	564	10	Q9XHC5
5	1789	63.6	569	10	Q9LYU8
6	1784	63.4	569	10	Q23152
7	1747.5	62.1	544	10	Q9FMU4
8	1737.5	61.8	544	10	Q23653
9	1737.5	61.8	544	10	Q9FY44
10	1675.5	59.6	559	10	Q95702
11	1570	55.8	354	10	Q9MAX0
12	635.5	24.7	470	16	Q8KG73
13	685	24.4	467	17	Q8RT30
14	667.5	23.6	479	17	Q8PX05
15	657.5	23.4	472	17	Q8TUD6
16	642.5	22.8	462	17	Q29558

ALIGNMENTS

RESULT 1
ID O8LPZ8 PRELIMINARY; PRT; 510 AA.
AC O8LPZ8;
DT 01-OCT-2002 (TREMBLrel. 22, Last created)
DT 01-OCT-2002 (TREMBLrel. 24, Last annotation update)
DE Monofunctional aspartokinase-like protein (EC 2.7.2.4) (Aspartate kinase).
DE OSUNBA0093F6.26.
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzae; Oryza.
OX NCBI_Taxid=39947;
RN [1]
RP SEQUENCE FROM N.A.
RT SRTAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, BAC clone:OSUNBA0093F6.26";
RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-ASPARTATE.
CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
CC EMBL, AP004332; BAB92806.1; -.
CC GRAMENE; Q8LPZ8; -.
DR GO: GO:0004072; P:aspartate kinase activity; IEA.
DR GO: GO:0016301; P:kinase activity; IEA.
DR GO: GO:0016740; P:transferase activity; IEA.
DR GO: GO:0008652; P:amino acid biosynthesis; IEA.
DR InterPro: IPR001048; Aa_kinase.
DR InterPro: IPR001341; Aspartate_kinase.
DR Pfam: PF00656; aa_kinase; 1.
DR TIGRFAMs: TIGR00657; asp_kinase; 1.
DR PROSITE: PS00324; ASPARTOKINASE; 1.
KW kinase; transferase.
SEQUENCE 510 AA; 54670 MW; 12C1603C1DF0D3EE CRC64;
SQ

17 607 21.6 473 16 Q9RUL9
18 605.5 21.5 818 16 Q8XAX1
19 590 21.0 461 16 Q8ZAS1
20 579.5 20.6 439 16 Q8A7Z9
21 569 20.2 449 16 Q8X5Y1
22 569 20.2 449 16 Q83IN8
23 566 20.1 449 16 Q8ZK15
24 566 20.1 449 16 Q8FB45
25 566 20.1 449 16 Q8ZIU8
26 550.5 19.6 450 16 Q9CM97
27 540.5 19.2 451 16 Q8EAC1
28 534 19.0 450 16 Q87J96
29 531 18.9 479 16 Q9KUM8
30 519 18.5 811 16 Q8A541
31 517.5 18.4 451 16 Q7VLP7
32 500.5 17.8 430 16 Q8DCU6
33 490.5 17.4 815 16 Q9CPD4
34 485.5 17.3 819 16 Q87SD0
35 483.5 17.2 825 16 Q8KPK3
36 477.5 17.0 820 16 Q8XAB4
37 477.5 17.0 820 16 Q7UDU5
38 477.5 17.0 834 16 Q83MT0
39 477.5 17.0 841 16 Q8FLD9
40 476.5 16.9 820 2 Q8RMX0
41 476.5 16.9 835 16 Q8PLH8
42 470.5 16.7 835 16 Q8P901
43 469.5 16.7 820 16 Q8ZS19
44 469.5 16.7 820 16 Q8ZSR7
45 468.5 16.7 454 16 Q8DR20

Q9RUL9 deinococcus
Q8XAX1 chlorobium
Q8ZAS1 yersinia pe
Q8A7Z9 bacteroides
Q8X5Y1 escherichia
Q83IN8 shigella fl
Q8ZK15 salmonella
Q8FB45 escherichia
Q8ZIU8 salmonella
Q9CM97 pasteurella
Q8EAC1 shewanella
Q87J96 vibrio para
Q9KUM8 vibrio chol
Q8A541 bacteroides
Q7VLP7 haemophilus
Q8DCU6 pasteurella
Q9CPD4 pasteurella
Q87SD0 vibrio para
Q8KPK3 vibrio chol
Q8XAB4 escherichia
Q7UDU5 shigella fl
Q83MT0 shigella fl
Q8FLD9 escherichia
Q8RMX0 escherichia
Q8PLH8 xanthomonas
Q8P901 xanthomonas
Q8ZS19 salmonella
Q8ZSR7 salmonella
Q8DR20 streptococc

Query Match 66.9%; Score 1881; DB 10; Length 510;
 Best Local Similarity 70.0%; Pred. No. 2,66-136;
 Matches 403; Conservative 29; Mismatches 56; Indels 88; Gaps 9;

QY 1 MAIPVRSAAAP--RLVPSIFPASG-----HYRGLACFGRTRGPRGARG 43
 DB 1 MAIALRLAAAPRLRLVSPAPPAPPAIGAGDGRGGGGAARIGVGRACRRRGGGRLE 60
 QY 44 LSMVADSTR-RAQADGGGCVGAPVLGSLGHEGLDQLSYMKRGSSVSAARMAE 102
 DB 61 VAAAAADDSARCRAKV--GAAAAAETLGGIGVGG--GDQLSYMKRGSSVSAARMAE 117
 QY 103 VAGLLTPPEERPVVLLSAMGKTNNLLAGEKAVGCGVHVSEIEEMNVKSLIKTYD 162
 DB 118 VAGLLTPPEERPVVLLSAMGKTNNLLAGEKAVGCGVHVSEIEEMNVKSLIKTYD 177
 QY 163 ELGLP-XICNTSLYELBOLLKGIAMMKELTPRTSDYLVSPGECSTRIFFSAVINKIRVKA 221
 DB 178 ELALPRSVIHTMLDELEQLKGIAMMKELTPRTSDYLVSPGECSTRIFFSAVINKIRVKA 237
 QY 222 RQYDAFDIGFITTDPRGNADILEATYPAVAKRLHGMWIDPAIPVYTGFLGKMGKSGAVT 281
 DB 238 RQYDAFDIGFITTDPRGNADILEATYPAVAKRLHGMWIDPAIPVYTGFLGKMGKSGAVT 297
 QY 282 TLGRGSGDLTATTIGKALGLREIQVWKVDGVLTCDPNIYPAKTYPYLTBEATLAYF 341
 DB 298 TLGRGSGDLTATTIGKALGLREIQVWKVDGVLTCDPNIYPAKTYPYLTBEATLAYF 357
 QY 342 GAGVLAHPSMRPARESDIPVRKNSYNPKAPGTLTRQDMDGVLVTSIVLSKNTML 401
 DB 358 GAGVLAHPSMRPARESDIPVRKNSYNPKAPGTLTRQDMDGVLVTSIVLSKNTML 415
 QY 402 DIVSTRMLQGYGLFARVSGICYIEDLCISVDCVATSEVSISLDPKIMSRRELIQOASE 461
 DB 416 DIVSTRMLQGYGLFARVSGICYIEDLCISVDCVATSEVSISLDPKIMSRRELIQOASE 471
 QY 462 LDHVELEKIAIVRLLOORATISLIGNEOSSILLEKTRVLRKSGVNVQMISQASKV 521
 DB 433 LDHVELEKIAIVRLLOORATISLIGNEOSSILLEKTRVLRKSGVNVQMISQASKV 521
 QY 522 NMSLIHDSDAKALVEALHQAFFEDVLSQVEAENL 557
 DB 472 NMSLIHDSDAKALVEALHQAFFEDVLSQVEAENL 507

RESULT 2
 Q85126 PRELIMINARY; PRT; 565 AA.
 ID Q85126 AC Q85126;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative aspartate kinase.
 GN OSJNB0015N08.5.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Erihartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA STRAIN=cv. Niponbare;
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
 RA Overton II L.L., Teitlin T., Kim M.M., Bera J.J., Jin S.S.,
 RA Faderon D.W., Tallon L.J., Koo H., Zisemann V., Heide J., Blunt S.,
 RA Vanaken S.S., Riedmuller S.B., Uteback T.T., Feldlyum T.V.,
 RA Yang Q.Q., Haas B.J., Sub B.B., Peterson J.J., Quackenbush J.,
 RA White O., Salzberg S.L., Frazer C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNB0015N08 genomic sequence";
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBD databases.
 DR GO: GO:0016597; F:amino acid binding; IEA.
 DR GO: GO:0004072; F:aspartate kinase activity; IEA.

DR GO: GO:0016301; F:kinase activity; IEA.
 DR GO: GO:0008652; P:amino acid biosynthesis; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR001048; Aa kinase.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aak_kinase; 1.
 DR Pfam: PF01842; ACT; 2.
 DR TIGRFAMs: TIGR00657; aep_kinase; 1.
 SQ SEQUENCE 565 AA; 60865 MW; 83103648C6FD7867 CRC64;

Query Match 66.8%; Score 1880; DB 10; Length 565;
 Best Local Similarity 70.7%; Pred. No. 3,66-136;
 Matches 396; Conservative 49; Mismatches 101; Indels 14; Gaps 8;

QY 1 MAIPVRSAAAP--RLVPSIFPASGAVRGLAC-FGRTRGPRGARGLSMVVADSTRRAKO 58
 DB 1 MAVALRFAAVARDPAPAAAPPRVGRBQVLAQAAAPRGRCRRRGLVVRGSGAAAV 60
 QY 59 ADGGDGVLAGVGLGMEGLGDLQSVMKFFGSSVSSAARMAEVAGLLTPPEERPVV 118
 DB 61 LNKD---AASVAAAAAASATG--FTVAMKFGSSVSAARMEVADLLSPFEETPVV 115
 QY 119 LSAMGKTNNLLAGEKAVGCGVHVSEIEEMNVKSLIKTYDEGLP-XICNTSLYEL 177
 DB 116 LSAMGKTNNLLAGEKAVGCGVHVSEIEEMNVKSLIKTYDEGLP-XICNTSLYEL 175
 QY 178 BOLLKGIAMMKELTPRTSDYLVSPGECSTRIFFSAVINKIRVKAQYDAFDIGFITTDE 237
 DB 176 BOLLKGIAMMKELTPRTSDYLVSPGECSTRIFFSAVINKIRVKAQYDAFDIGFITTDE 235
 QY 238 GNAIDLEATYPAVAKRLHGMWIDPAIPVYTGFLGKMGKSGAVTTLGRGSDLTATTIGK 297
 DB 236 GNAIDLEATYPAVAKRLHGMWIDPAIPVYTGFLGKMGKSGAVTTLGRGSDLTATTIGK 295
 QY 298 ALGLREIQVWKVDGVLTCDPNIYPAKTYPYLTBEATLAYFGAQLHPQMRPARE 357
 DB 296 ALGLREIQVWKVDGVLTCDPNIYPAKTYPYLTBEATLAYFGAQLHPQMRPARE 355
 QY 358 DIPVRKNSYNPKAPGTLTRQDMDGVLVTSIVLSKNTMLDIVSTRMLQGYGLAR 417
 DB 356 DIPVRKNSYNPKAPGTLTRQDMDGVLVTSIVLSKNTMLDIVSTRMLQGYGLAR 413
 QY 418 VSGICYIEDLCISVDCVATSEVSISLDPKIMSRRELIQOASELDHVELEKIAIVRL 477
 DB 414 VSGI--FEDLGISVDCVATSEVSISLDPKIMSRRELIQ--LDHVELEKIAIVRL 469
 QY 478 LQORATISLIGNEOSSILLEKTRVLRKSGVNVQMISQASKVNMGLYHDSAKALVE 537
 DB 470 LQORATISLIGNEOSSILLEKTRVLRKSGVNVQMISQASKVNMGLYHDSAKALVE 529
 QY 538 ALHQAFFEDVLSQVEAENL 557
 DB 530 ALHQAFFEDVLSQVEAENL 549

RESULT 3
 Q88YL8 PRELIMINARY; PRT; 581 AA.
 ID Q88YL8 AC Q88YL8;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Putative aspartate kinase (EC 2.7.2.4) (Aspartokinase).
 GN OSJNB00052012.2.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Erihartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RC SEQUENCE FROM N.A.

RC STRAIN-cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT Oryza sativa nipponbare (GA) genomic DNA, chromosome 1, BAC
 RT clone:OSUNB0052012.1;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
 CC ASPARTATE.
 CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
 DR EMBL: AF004330; BAB90744.1; -.
 DR Gramene: Q8RY18; -.
 DR GO: GO:0004072; F:aspartate kinase activity; IEA.
 DR GO: GO:0016501; F:kinase activity; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:0008652; P:amino acid biosynthesis; IEA.
 DR InterPro: IPR001048; Aa kinase.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aakinese; 1.
 DR TIGRfam: TIGR00657; asp_kinases; 1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 KW Kinase; Transferase.
 SQ SEQUENCE 581 AA; 62638 MW; AD8B0811EB906AAB CRC64;

Query Match 66.4%; Score 1868.5; DB 10; Length 581;
 Best Local Similarity 66.0%; Pred. No. 2.9e-135;
 Matches 409; Conservative 34; Mismatches 72; Indels 105; Gaps 11;

QY 1 MAIPVSAAP-RLVPSIIPASSG-----HRLGACPTTGTGPRARG 43
 DB 1 MAIALRLAALRLRLVSPAPPAICAGDGRGGGAARIGVLGRACRRRRGGGGTLE 60
 QY 44 LSNVVDNSR-BAKQADGGDVLGAPVLGHEGLDQLSVNKKFGSSSVSAAAPAE 102
 DB 61 VAAAADDSRCRAKAY-GAAAAAETGGGIVGG-GQLSVNKKFGSSSVSAAAPAE 117
 QY 103 VAGLLTFPEERPVVLSAMGKTTNNLLAGEKAVGCVIHYSEIEMNWKSLIKTV 162
 DB 118 VAGLLTFPEERPVVLSAMGKTTNNLLAGEKAVGCVIHYSEIEMNWKSLIKTV 177
 QY 163 ELGLP-XICNTSLYEQLKGIAMKELTPRTDVLVSFGCSMTRIFSAVLNIRKA 221
 DB 178 ELALPSVHTMDELQQLKGIAMKELTLRTDVLVSFGCSMTRIFSAVLNIRKA 237
 QY 222 RQYDAFDIGFTTDEFGNADILEATYPAVAKRLHGMIDDPALPVYTGFLGKMGSAVT 281
 DB 238 RQYDAFDIGFTTDDGNADILEATYPAVAKRLHGMIDDPALPVYTGFLGKMGSAVT 297
 QY 282 TLGRGSDLTATTITGALGLREIQWKKVDVGLTCDPNITYPRAKIVPYLTPEAEIAYF 341
 DB 298 TLGRGSDLTATTITGALGLREIQWKKVDVGLTCDPNITYPRAKIVPYLTPEAEIAYF 357
 QY 342 GAQVLYHQSRPARSGDIPRVKSNPNRAPGTLIRCRDMXGVLVLTSLKSNVTML 401
 DB 358 GAQVLYHQSRPARSGDIPRVKSNPNRAPGTLIRCRDMX--VVLTSILKSNVTML 415
 QY 402 DIVSTRMLGQYGLARVSGICYEDLCISVDCAVSEVSUSLDSKIMSELIQAS- 460
 DB 416 DIVSTRMLGQYGLARVSGICYEDLCISVDCAVSEVSUSLDSKIMSELIQAS- 443
 QY 461 -----ELDHVELEKIALVRLLOQRATIS 485
 DB 444 PIYKQNVVILVAGISDQSTTFQVLFALPLPFGDLHVELEKIALVRLLOQRATIS 503
 QY 486 LIGNVQSSLILEKTGVLRK-----SGVNVQMI SQASKVNSLIVHSDAKALVE 537
 DB 504 LIGNVRRSLILEKDSNHTHTHTYTSGLGLSSVEKGECPDILA-----SEAKQCIK 558
 QY 538 ALHQAFEDDVLISQVEANL 557
 DB 559 ALHQAFEDDVLIEVEEAL 578

ID 09XHC5 PRELIMINARY; PRT; 564 AA.
 AC 09XHC5;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Precursor monofunctional aspartokinase (EC 2.7.2.4) (Aspartate
 DE kinase).
 OS Glutamine max (soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosidia I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_Taxid:3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Century;
 RA Beau B.D., Frankard V., Jacobs M., Matthews B.F.;
 RT "Isolation and characterization of a cDNA clone encoding a
 RT monofunctional aspartokinase."
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
 CC ASPARTATE.
 CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
 DR EMBL: AF135862; BAD41796.1; -.
 DR GO: GO:0016597; P:amino acid binding; IEA.
 DR GO: GO:0004072; F:aspartate kinase activity; IEA.
 DR GO: GO:0016301; F:kinase activity; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:0008652; P:amino acid biosynthesis; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR001048; Aa kinase.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aakinese; 1.
 DR Pfam: PF01842; ACT; 2.
 DR TIGRfam: TIGR00657; asp_kinases; 1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 KW Kinase; Transferase.
 SQ SEQUENCE 564 AA; 61285 MW; 5835C4F855FC6157 CRC64;

Query Match 65.2%; Score 1835; DB 10; Length 564;
 Best Local Similarity 70.6%; Pred. No. 1.1e-132;
 Matches 389; Conservative 42; Mismatches 80; Indels 40; Gaps 7;

QY 37 GPRGARGLSVYVADSTRRAKQADGDEVLGAPV----- 71
 DB 10 GVQKLAIVMSVRSRLHCKSQI--GPAALGAPVCARRWGNRAVSYTCKASTDVE 67
 QY 72 -----GGLGMEGLDQLSVNKKFGSSSVSAAAPAEVAGILTFPEERPVVLSAMGKT 125
 DB 68 KNATENGWVSSEG-ETSFYCVNKKFGSSSVASADRKKKEVATILSPFERPIVLSAMGKT 126
 QY 126 TNNLLAGEKAVGCGVIVHYSEIEMNWKSLIKTVDELGLP-XICNTSLYEQLKGI 184
 DB 127 TNNLLAGEKAVGCGVIVHYSEIEMNWKSLIKTVDELGLP-XICNTSLYEQLKGI 186
 QY 185 AMKELTPRTDVLVSFGCSMTRIFSAVLNIRKAQYAPDGLFTTDEFGNADILE 244
 DB 187 AMKELTPRTDVLVSFGCSMTRIFSAVLNIRKAQYAPDGLFTTDEFGNADILE 246
 QY 245 ATYPAVAKRLHGMIDDPALPVYTGFLGKMGSAVTLTGRGSDLTATTITGALGLREI 304
 DB 247 ATYPAVAKRLHGMIDDPALPVYTGFLGKMGSAVTLTGRGSDLTATTITGALGLREI 306
 QY 305 QWKKVDVGLTCDPNITYPRAKIVPYLTPEAEIAYFPAQVLYHQSRPARSGDIPRVK 364
 DB 307 QWKKVDVGLTCDPNITYPRAKIVPYLTPEAEIAYFPAQVLYHQSRPARSGDIPRVK 366
 QY 365 NSYNPAPGTLIRCRDMXGVLVLTSLKSNVTMLDIVSTRMLGQYGLARVSGICY 424
 DB 367 NSYNPAPGTLIRKADMSK--AVLTSILKSNVTMLDIVSTRMLGQYGLARVSGICY 422
 QY 425 EDLCISVDCAVSEVSUSLDSKIMSELIQASDELHVELEKIALVRLLOQRATIS 484

DB 423 EELGISVDVATSEVSVSLTLPDKLMSRELIIQASELDHYVELEKIAVNLQNRSII 462
 QY 485 SLIGNEQSSLLIEKTRVLRKSGVNVQMSIQASKNMSLITVDSKALVELHQAFF 544
 DB 483 SLIGNVRSLLIERLSRVLRTLVQVQMSIQASKNMSLIVNDSKALVELHQAFF 542
 QY 545 EDDVLSQVEAE 555
 DB 543 ESE-LSLELMD 552

RESULT 5
 Q9LYU8 PRELIMINARY; PRT; 569 AA.
 ID Q9LYU8
 AC Q9LYU8
 DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 DE Aspartate kinase (EC 2.7.2.4) (AT5G13280/T31B5_100)
 DE (Aspartokinase).
 GN T31B5.100.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Hlbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
 RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai U., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.T., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Becker J.R.;
 RT "Arabidopsis cDNA clones";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Shim P., Bowser L., Carninci P.,
 RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
 RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai U., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Palm C.J., Quach H.T., Sakurai T., Satou M., Seki M.,
 RA Southwick A., Tang C.C., Toriumi M., Wallender B.K., Wong C., Wu H.C.,
 RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
 RA Becker J.R.;
 RT "Arabidopsis ORF clones";
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 CC -|- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
 CC ASPARTATE.
 CC -|- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
 CC EMBL: AL163491; CAB86635.1; -
 CC EMBL: AY057674; AAL15305.1; -
 CC EMBL: BT000493; AAN18062.1; -
 CC PIR: T48575; T48575.
 CC GO: GO:001697; P.amino acid binding; IEA.
 CC GO: GO:0004072; Aspartate kinase activity; IEA.
 CC GO: GO:0016501; F.kinase activity; IEA.
 CC GO: GO:0016740; F.transferase activity; IEA.
 CC GO: GO:0008652; P.amino acid biosynthesis; IEA.
 CC GO: GO:0008152; P.metabolism; IEA.
 CC InterPro: IPR001046; Aa_kinase.

DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aak_kinase; 1.
 DR TIGRfam: TIGR01842; ACT; 2.
 DR TIGRfam: TIGR00657; asp_kinases; 1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 KW kinase; transferase.
 SQ SEQUENCE 569 AA; 62298 MW; F66A3F4E84DC429 CRC64;

Query Match 63.6%; Score 1789; DB 10; Length 569;
 Best Local Similarity 71.8%; Pred. No. 3.8e-129;
 Matches 369; Conservative 54; Mismatches 79; Indels 12; Gaps 6;

QY 44 LSMVADSTRRAQADGGDGVLAGAPVLAGLGMGLGD---QLSVMKFGSSSVSSAAR 99
 DB 44 LSLPFGDSSIRKYSGSGSRNIVRA-VLEEKTEALTEVDEKGITCVMKFGSSVASAR 102
 QY 100 MAEVAGLLTTPPERPVVYLSAMKTTNNLLGKRAVGGCVIHVEIEEMNVKSLHK 159
 DB 103 MKEVADLITPPESPVIVLSAMKTTNNLLGKRAVGGCVNASEIEELITIKELHR 162
 QY 160 TVDELGL-PIKNTSLYELEQLKGIAMKELPTISDYLVSEGCMSIRIFSAVUNKR 218
 DB 163 TKELNIDPSVILVLELEQLKGIAMKELTRTDVLSFGECISTRFPAVYNTIG 222
 QY 219 VKARQYDAFDIGFTITDFEFGADILEATYPAVAKRLHGMVQPAIPVYTFGKWKSG 278
 DB 223 VKARQYDAFEIGFTITDFEFGADILEATYPAVAKRLHGMVQPAIPVYTFGKWKSG 282
 QY 279 AVTTGGRGSDLTATTGKALGLREIQVMDVDGVLTCDNINYPHAKTVPYLTFEBATEL 338
 DB 283 AVTTGGRGSDLTATTGKALGLREIQVMDVDGVLTCDNINYPHAKTVPYLTFEBATEL 342
 QY 339 AVEGAQVILHPSMKPARBODIPVAKSYNPKAPGTLITQRMDXGLVLTSTYKSNV 398
 DB 343 AVEGAQVILHPSMKPARBODIPVAKSYNPKAPGTLITQRMDXGLVLTSTYKSNV 400
 QY 399 TMLDVTSTRMIGOVGLARVSGICIEDLCISVDQVATSEVSVSLSDPKMSRELIIQ 458
 DB 401 TMLDIASIRMGVGFIAKVRST--FEELGISVDVATSEVSVSLTLPDKLMSRELIIQ 458
 QY 459 ASELHDHYVELEKIAVRLIQORAIISLIGNEQSSLLIEKTRVLRKSGVNVQMSIQ 518
 DB 459 --ELHDHYVELEKIAVNLKGRAIISLIGNEQSSLLIERAPHVYTKGVNVQMSIQ 516
 QY 519 SKVNMSTLVHSDAKALVELHQAFFEDVLSQV 552
 DB 517 SKVNISFTVNEAEAGCVQALHKSFFESGDLSEL 550

RESULT 6
 ID Q23152 PRELIMINARY; PRT; 569 AA.
 AC Q23152
 DT 01-JAN-1998 (TrEMBLrel. 05, last sequence update)
 DT 01-JAN-1998 (TrEMBLrel. 05, last annotation update)
 DE Aspartate kinase precursor (EC 2.7.2.4) (Aspartokinase).
 GN AK-LYS1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97351557; PubMed=9207839;
 RT Frankard V., Vauterin M., Jacobs M.;
 RT "Molecular characterisation of an Arabidopsis thaliana cDNA coding for
 a monofunctional aspartate kinase";
 RT Plant Mol. Biol. 34:233-242 (1997).
 CC -|- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
 CC ASPARTATE.

CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
 DR EMBL; X98873; CA67376.1; -;
 DR GO; GO:0016597; F:amino acid binding; IEA.
 DR GO; GO:0004072; F:aspartate kinase activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR Pfam; PF00696; aak_kinase; 1.
 DR Pfam; PF01842; ACT; 2.
 DR TIGRFAMs; TIGR00657; asp_kinases; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 KM Kinase; Signal; Transferase.
 FT SIGNAL 1 86
 FT CHAIN 87 569 ASPARTATE KINASE.
 SQ SEQUENCE 569 AA; 62272 MW; CD5323120D94DD33 CRC64;
 Query Match 63.4%; Score 1784; DB 10; Length 569;
 Best Local Similarity 71.6%; Pred. No. 9, 2e-129;
 Matches 368; Conservative 55; Mismatches 79; Indels 12; Gaps 6;

QY 44 LSNVVDSTRRKQADGQGVGAPVLGSLGMEGLD---QLSVYMKFGSSVSVAAR 99
 DB 44 LSLPTGGSSIRKVSQSGSNIVRA-VLEKKTAEITVEDEKGITVWKFQSSVSAAR 102
 QY 100 MAEYAGLITFPERPVPVVLASAKTNNLLAGEKAVGCVIHVSEIEEMVXSLHX 159
 DB 103 MKEVADILTFPEESPVTIVASAKTNNLLAGEKAVGCVHSEIEELSIKELHIR 162
 QY 160 TVDEELG-EXICNTSLVELFQLKGTAMKELTPRSDIVYSGEOMSRISAYINKR 218
 DB 163 TVKELNTDPSVILTYLEELQOLKGTAMKELTRFDIVYSGELSTRIPAAVYNTIG 222
 QY 219 VKARQYDAFDIGITTFDEFGNADILEATYPAVAKRLHGMIDPAIPVVTGFLGKQWSG 278
 DB 223 VKARQYDAFEIGITTFDDFTNGDILEATYPAVAKRLYDDMDHPAIPVITGFLGKQWKG 282
 QY 279 AVTTLGRGSDLTATTIGKALGIREIOWKDVGVLTCPDNITYPAKTYPIYTFEATL 338
 DB 283 AVTTLGRGSDLTATTIGKALGIREIOWKDVGVLTCPDNITYPAKTYPIYTFEATL 342
 QY 339 AYFGAQLHPQSMRPAEEDIPVRKNSYNPKAPGTLITRQMDGLVLTSLYKSNV 398
 DB 343 AYFGAQLHPQSMRPAEEDIPVRKNSYNPKAPGTLITRQMDGLVLTSLYKSNV 400
 QY 399 TMDIVSTRMLGOYGLFARVSGICYIEDLCISVDCVATSEVSVSLDPSKIMSRLEIQ 458
 DB 401 TMDIVSTRMLGOYGLFARVSGICYIEDLCISVDCVATSEVSVSLDPSKIMSRLEIQ 458
 QY 459 ASELDHYVEELEKAIYRLIQARISLIGNVQSSSLILEKTRVLRKSNVNMVMSOGA 518
 DB 459 ASELDHYVEELEKAIYRLIQARISLIGNVQSSSLILEKTRVLRKSNVNMVMSOGA 516
 QY 519 SKVMSLIIVHSDAKALVEALHQAFFEDVLSQV 552
 DB 517 SKVMSLIIVHSDAKALVEALHQAFFEDVLSQV 550

RESULT 7
 Q9FMU4 PRELIMINARY; PRT; 544 AA.
 AC Q9FMU4;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Lysine-sensitive aspartate kinase (EC 2.7.2.4)
 OS (Aspartokinase)
 OS Arabidopsis thaliana (Mouse-ear cress)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=98162728; PubMed=9501997;
 RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
 RT Sequence features of the regions of 1,191,918 bp covered by seventeen
 RT physically assigned P1 clones."
 RI DNA Res. 4:401-414 (1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSHO-L-
 CC ASPARTATE.
 CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
 DR EMBL; AB007650; BAB08285.1; -;
 DR GO; GO:0016597; F:amino acid binding; IEA.
 DR GO; GO:0004072; F:aspartate kinase activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR Pfam; PF00696; aak_kinase; 1.
 DR Pfam; PF01842; ACT; 2.
 DR TIGRFAMs; TIGR00657; asp_kinases; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 KM Kinase; Transferase.
 SQ SEQUENCE 544 AA; 59604 MW; 7DBCFDC1138645AC CRC64;
 Query Match 62.1%; Score 1747.5; DB 10; Length 544;
 Best Local Similarity 74.5%; Pred. No. 5, 6e-126;
 Matches 351; Conservative 50; Mismatches 65; Indels 5; Gaps 3;

QY 78 GLGDQSLVYMKFGSSVSAAAEVAGLITFPERPVPVVLASAKTNNLLAGEKAV 137
 DB 77 GTGKEELCVWKFQSSVESAERKEVANIILSPDEREIVIVASAKTNNLLAGEKAV 136
 QY 138 GCGVIVHSEIEEMVXSLHXIKTVDELG-EXICNTSLVELFQLKGTAMKELTPRSD 196
 DB 137 TCGVIVHSEIEELQOLKGTAMKELTRFDIVYSGELSTRIPAAVYNTIG 196
 QY 197 YLVSPGECSTRIFSAYLNKIRKARQYDAFDIGITTFDEFGNADILEATYPAVAKRLHG 256
 DB 197 YLVSPGECSTRIFSAYLNKIRKARQYDAFDIGITTFDDFTNGDILEATYPAVAKRLVG 256
 QY 257 DWIODPAIPVVTGFLGKQWSGAVTTLGRGSDLTATTIGKALGIREIOWKDVGVLTG 316
 DB 257 DWKRENAVVPVVTGFLGKQWSGAVTTLGRGSDLTATTIGKALGIREIOWKDVGVLTG 316
 QY 317 DPNITYPAKTYPIYTFEATLSELDHYVEELEKAIYRLIQARISLIGNVQSSSLILEK 376
 DB 317 DPNITYPAKTYPIYTFEATLSELDHYVEELEKAIYRLIQARISLIGNVQSSSLILEK 376
 QY 377 TROQMDXGLVLTSLYKSNVMTMDIVSTRMLGOYGLFARVSGICYIEDLCISVDCVAT 436
 DB 377 TROQMDXGLVLTSLYKSNVMTMDIVSTRMLGOYGLFARVSGICYIEDLCISVDCVAT 432
 QY 437 SEVSVSLDPSKIMSRLEIQARISLIGNVQSSSLILEKTRVLRKSNVNMVMSOGA 496
 DB 437 SEVSVSLDPSKIMSRLEIQARISLIGNVQSSSLILEKTRVLRKSNVNMVMSOGA 492
 QY 497 LEKTRVLRKSNVNMVMSOGASKNVMSLIIVHSDAKALVEALHQAFFEDD 547
 DB 493 LEKTRVLRKSNVNMVMSOGASKNVMSLIIVHSDAKALVEALHQAFFEDD 543

RESULT 8
 O23653 PRELIMINARY; PRT; 544 AA.
 ID O23653
 AC O23653;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Lysine-sensitive aspartate kinase (EC 2.7.2.4)
 DE (Aspartokinase).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 NC NCB1 Taxid=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97351562; PubMed=9207844;
 RA Tang G., Zhu-Shimoni J.X., Amir R., Zehori I.B., Galili G.,
 RT "Cloning and expression of an Arabidopsis thaliana cDNA encoding a
 RT monofunctional aspartate kinase homologous to the lysine-sensitive
 RT enzyme of Escherichia coli.";
 RL Plant Mol. Biol. 34:287-294(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
 CC ASPARTATE.
 CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
 DR EMBL: U62020; AAB63104.1; -.
 DR GO: GO:0016597; Amino acid binding; IEA.
 DR GO: GO:0004072; Aspartate kinase activity; IEA.
 DR GO: GO:0016301; P.kinase activity; IEA.
 DR GO: GO:0016740; P.kinase activity; IEA.
 DR GO: GO:0006552; Amino acid biosynthesis; IEA.
 DR GO: GO:0008152; P.metabolism; IEA.
 DR InterPro: IPR001048; Aa_kinase.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aak_kinase; 1.
 DR Pfam: PF01842; ACT; 2.
 DR TIGRFAMs: TIGR00657; asp_kinases; 1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 DR KINASE; Transferase.
 KM KINASE; 544 AA; 59506 MW; 2C204BCEB23DD80 CRC64;
 SQ SEQUENCE

Query Match 61.8%; Score 1737.5; DB 10; Length 544;
 Best Local Similarity 74.1%; Pred. No. 3.3e-125;
 Matches 349; Conservative 50; Mismatches 67; Indels 5; Gaps 3;

QY 78 GLGDQSLVYMKFGSSVSSAARMAEYAGLILTPFERPPVVLASMGKTTNNLLAGEKAV 137
 DB 77 GTGKELTCVMKFGSSVSSAARMEKAVANLIFSPDERPVILASMGKTNKLLKAGEKAV 136
 QY 138 GCGVIVHSEIEEMNWKSLIKTYDELGL-PXICNTSYELBQLKGIAMMKELTPRTSD 196
 DB 137 TCGVTVGSEIEELSLIKELHRLTAHELGVETTVIEKLEGLHQLKGISMKKELTLRTD 196
 QY 197 YLVFGECHSTRIFSAVYLNKIRVAKQYDAFIDGFTTDFEGNADILEATYPAVAKLHG 256
 DB 197 YLVFGECHSTRIFSAVYLNKIRVAKQYDAFIDGFTTDFEGNADILEATYPAVAKLHG 256
 QY 257 DNIQDAIPVVTGFLGKMKSGAVTTLGRGSDLTATTIGKALGRLIOWKMDVGVLTLC 316
 DB 257 DMSKENAVPVYTGFLGKMKSCAITTLGRGSDLTATTIGKALGRLIOWKMDVGVLTLC 316
 QY 317 DENIYPHAKTVPYLTFEATELAYFGAQLVLPQSMRPARBEDIIVRVKNSYNPAFGTLI 376
 DB 317 DNIYPHAKTVPYLTFEATELAYFGAQLVLPQSMRPARBEDIIVRVKNSYNPAFGTLI 376
 QY 377 TRORDMXGLVLTSLVLSKSNVTMLDIVSTRLMGQYGLFARVSGICYEDLCISVDVAT 436
 DB 377 TRSRDMK--AVLTSLVLSKSNVTMLDIVSTRLMGQYGLFARVSGICYEDLCISVDVAT 432
 QY 437 SEVSVSLSLDPKSRRELIOQASGLDHVVELEKIAIVRLIOQRATISLIGNVQSSLI 496
 DB 433 SEVSISLTLDPKALRGRELIOQVNELDNVELEKIAVYKLIQRRSIIISLIGNVQSSLI 492
 QY 497 LEKTVGLRSGVNVOMISOGAKVMSLIIVHSDPAKALVEALHQAFFEDD 547
 DB 493 LEKTVGLRSGVNVOMISOGAKVMSLIIVHSDPAKALVEALHQAFFEDD 543

RESULT 9
 ID 09FY44 PRELIMINARY; PRT; 544 AA.
 AC 09FY44;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Aspartate kinase precursor (EC 2.7.2.4) (Aspartokinase).
 GN AK-LYS2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 NC NCB1 Taxid=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=Leaf;
 RA Frankard V.M.S., Vauterin M., Jacobs M.;
 RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
 CC ASPARTATE.
 CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
 DR EMBL: Y16255; CAC06395.1; -.
 DR GO: GO:0016597; Amino acid binding; IEA.
 DR GO: GO:0004072; Aspartate kinase activity; IEA.
 DR GO: GO:0016301; P.kinase activity; IEA.
 DR GO: GO:0016740; P.kinase activity; IEA.
 DR GO: GO:0006552; Amino acid biosynthesis; IEA.
 DR GO: GO:0008152; P.metabolism; IEA.
 DR InterPro: IPR001048; Aa_kinase.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aak_kinase; 1.
 DR Pfam: PF01842; ACT; 2.
 DR TIGRFAMs: TIGR00657; asp_kinases; 1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 DR KINASE; Signal; Transferase.
 FT SIGNAL 1
 SQ SEQUENCE 544 AA; 59578 MW; 3308371224C309A6 CRC64;

Query Match 61.8%; Score 1737.5; DB 10; Length 544;
 Best Local Similarity 74.1%; Pred. No. 3.3e-125;
 Matches 349; Conservative 50; Mismatches 67; Indels 5; Gaps 3;

QY 78 GLGDQSLVYMKFGSSVSSAARMAEYAGLILTPFERPPVVLASMGKTTNNLLAGEKAV 137
 DB 77 GTGKELTCVMKFGSSVSSAARMEKAVANLIFSPDERPVILASMGKTNKLLKAGEKAV 136
 QY 138 GCGVIVHSEIEEMNWKSLIKTYDELGL-PXICNTSYELBQLKGIAMMKELTPRTSD 196
 DB 137 TCGVTVGSEIEELSLIKELHRLTAHELGVETTVIEKLEGLHQLKGISMKKELTLRTD 196
 QY 197 YLVFGECHSTRIFSAVYLNKIRVAKQYDAFIDGFTTDFEGNADILEATYPAVAKLHG 256
 DB 197 YLVFGECHSTRIFSAVYLNKIRVAKQYDAFIDGFTTDFEGNADILEATYPAVAKLHG 256
 QY 257 DNIQDAIPVVTGFLGKMKSGAVTTLGRGSDLTATTIGKALGRLIOWKMDVGVLTLC 316
 DB 257 DMSKENAVPVYTGFLGKMKSCAITTLGRGSDLTATTIGKALGRLIOWKMDVGVLTLC 316
 QY 317 DENIYPHAKTVPYLTFEATELAYFGAQLVLPQSMRPARBEDIIVRVKNSYNPAFGTLI 376
 DB 317 DNIYPHAKTVPYLTFEATELAYFGAQLVLPQSMRPARBEDIIVRVKNSYNPAFGTLI 376
 QY 377 TRORDMXGLVLTSLVLSKSNVTMLDIVSTRLMGQYGLFARVSGICYEDLCISVDVAT 436
 DB 377 TRSRDMK--AVLTSLVLSKSNVTMLDIVSTRLMGQYGLFARVSGICYEDLCISVDVAT 432
 QY 437 SEVSVSLSLDPKSRRELIOQASGLDHVVELEKIAIVRLIOQRATISLIGNVQSSLI 496
 DB 433 SEVSISLTLDPKALRGRELIOQVNELDNVELEKIAVYKLIQRRSIIISLIGNVQSSLI 492

QY 497 LEKTRVLRKSGVNVQWISQASAKVNNSLIVHSDAKALVEALHQAFFEDD 547
Db 493 LEKTRVLRKSGVNVQWISQASAKVNNSLIVHSDAKALVEALHQAFFEDD 543

RESULT 10

Q9S702 PRELIMINARY; PRT; 559 AA.
AC Q9S702
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Putative aspartate kinase (EC 2.7.2.4) (Aspartokinase).
GN P109.20 OR F2807.35.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;

SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Bentto M., Creasy T.H., Haas B.,
RA Rensing C.M., Koo H., Fujii C.Y., Uterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RA "Arabidopsis thaliana chromosome III BAC P109 genomic sequence";
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 0:0-0(2002).

SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
CC ASPARTATE.

CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.

DR EMBL; AC011664; AAFA1833.1; -;
DR EMBL; AC010797; AAF03452.1; -;
DR EMBL; AY088366; AAM65905.1; -;
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:0004072; F:aspartate kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008652; P:metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR001341; Aspartate_kinase.
DR Pfam; PF00696; aak_kinase; 1.
DR TIGRfam; TIGR00657; asp_kinases; 1.
DR PROSITE; PS00324; ASPARTOKINASE; 1.
KW Kinase; Transferase.
SQ SEQUENCE 559 AA; 6115 MW; FRC84A0B814F349 CRC64;

Query Match 59.6%; Score 1675.5; DB 10; Length 559;
Best Local Similarity 73.5%; Pred. No. 2, Le-120;
Matches 347; Conservative 47; Mismatches 71; Indels 7; Gaps 4;

QY 82 QLSVVRKFGSSVSAAARMAEVLITPPEERPVVLSMGKTNLLAGKAVGCCV 141
Db 82 KLTGKFGSSVSAAARMAEVLITPPEERPVVLSMGKTNLLAGKAVGCCV 141
QY 142 IHVSEIEMNNVKSLLHKTVDLGL-PIXCNTSLVLELLKGIAMKKEITPTSTYLV 200

Db 142 TNDVTIELSLYIEHLIRTAHBLGVETAVIAHBLGELQKGVAMKEILTRSDYLV 201
QY 201 FGECMSTRIRFSAYLANKIRVAKQYDAFDIGFTTDEFGNADILEATYPAVARLHGMIQ 260
Db 202 FGECMSTRIRFSAYLANKIRVAKQYDAFDIGFTTDEFGNADILEATYPAVARLHGMIQ 261
QY 261 DPALPVVTLGKSGKSGAVTTLGRGSDLTATTTGKALGRIEIVMDVGVLCBPN 320
Db 262 ENALPVVTLGKSGKSGAVTTLGRGSDLTATTTGKALGRIEIVMDVGVLCBPN 321
QY 321 YPAKTVPYLTPEEATLAVFGAQLVHPSMRPAEGDIPVAKSYNPKAFGLITROR 380
Db 322 YCAQAPPHITFDEAELAVFGAQLVHPSMRPAEGDIPVAKSYNPKAFGLITROR 381
QY 381 DMDXGLVLTSTYLSKSNVMTDIVSRMLGQYGFARVSGIYIIDLCLSDCVATSEVS 440
Db 382 DMKR-AVLISYLSKSNVMTDIVSRMLGQYGFARVSGIYIIDLCLSDCVATSEVS 437
QY 441 VSVSLDPSKISRELLIQASELDHVELEKTAIVLLQORAILSLIGVQSSLLILEKT 500
Db 438 ISLTDPKSCSRELLQH--ELDQVEELEKTAIVVNLHRSIISLIGVQSSFTLEKG 495
QY 501 GRVLRKSGVNVQWISQASAKVNNSLIVHSDAKALVEALHQAFFEDDVLQY 552
Db 496 FRVLRKSGVNVQWISQASAKVNNSLIVHSDAKALVEALHQAFFEDDVLQY 547

RESULT 11

Q9MAXO PRELIMINARY; PRT; 354 AA.
AC Q9MAXO
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Aspartate kinase (EC 2.7.2.4) (Aspartokinase) (Fragment).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxID=4530;

SEQUENCE FROM N.A.

RA Kiyoita S.;
RT "Lysine sensitive aspartate kinase from rice";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
CC ASPARTATE.

CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.

DR EMBL; AB042521; BA95630.1; -;
DR Gramene; Q9MAXO; -;
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:0004072; F:aspartate kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008652; P:metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR001341; Aspartate_kinase.
DR Pfam; PF00696; aak_kinase; 1.
DR TIGRfam; TIGR00657; asp_kinases; 1.
KW Kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 354 AA; 3885 MW; 88B668815627A040 CRC64;

Query Match 55.8%; Score 1570; DB 10; Length 354;
Best Local Similarity 88.2%; Pred. No. 1, 4e-112;
Matches 315; Conservative 18; Mismatches 18; Indels 5; Gaps 3;

QY 201 FGECMSTRIRFSAYLANKIRVAKQYDAFDIGFTTDEFGNADILEATYPAVARLHGMIQ 260
Db 1 FGECMSTRIRFSAYLANKIRVAKQYDAFDIGFTTDEFGNADILEATYPAVARLHGMIQ 260

QY 261 DPAPVVTGFLGKMGKSGAVTTTIGRGSDLTATTIGALGLREIQVMKDVGVLTCDPNI 320
 DB 61 DPAPVVTGFLGKMGKSGAVTTTIGRGSDLTATTIGALGLREIQVMKDVGVLTCDPNI 120
 QY 321 YPAKATVPVLTPEEATLAVFGAQLVHPQSMAPAREGDI PVRVKSYNPKAPGTLITPQR 360
 DB 121 YPAATVPVLTPEEATLAVFGAQLVHPQSMAPAREGDI PVRVKSYNPKAPGTLITPQR 180
 QY 381 DMDGVLVLTISYLVKSNVTMLDIVSTRMLGQVFLARVSGICIEDLCISVDCVATSEVS 440
 DB 181 EMOK--VLTISYLVKSNVTMLDIVSTRMLGQVFLARVSGICIEDLCISVDCVATSEVS 236
 QY 441 VVSVDPSKTSWRELLIQASLHDHVEELEKTAIVYLLQQRAIISLIGNVDSGLIEKT 500
 DB 237 VVSVDPSKTSWRELLIQ--ELDHVEGLEKTAIVYLLQQRAIISLIGNVDSGLIEKTA 294
 QY 501 GRVLRKSGVNVQMIISQASRVNMSLIVHSDAKALVHALQAFFEDVLSQVEAENT 557
 DB 295 FGVLKRSQVNVQMIISQASRVNMSLIVHSDAKALVHALQAFFEDVLSQVEAENT 351
 RESULT 12
 Q8KG73 PRELIMINARY; PRT; 470 AA.
 AC Q8KG73;
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Aspartokinase.
 GN LYSC OR CT0095.
 OS Chlorobium tepidum.
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
 OC Chlorobium.
 NCBI_TaxID=1097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TLS / ATCC 49652 / DSM 12025;
 RX MEDLINE=22103685; PubMed=1203901;
 RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
 RA Dodson R.J., Deboy R., Gwin M.L., Nelson W.C., Haft D.H.,
 RA Hickey E.K., Peterson J.D., Durkin A.S., Kolony J.L., Yang F.,
 RA Holt I., Umayar L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
 RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
 RA Vamathevan J., Kouri H., Bryant D.A., Fraser C.M.,
 RA "The complete genome sequence of Chlorobium tepidum TIS, a
 RT photosynthetic, anaerobic, green-sulfur bacterium";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
 DR EMBL; AE012788; AAM71343.1;
 DR TIGR; CT0095;
 DR GO; GO:0016597; F:amino acid binding; IEA.
 DR GO; GO:0004072; F:aspartate kinase activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001046; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR InterPro; IPR000847; HTH_LysR.
 DR Pfam; PF00696; aakinese; 1.
 DR Pfam; PF01842; ACT; 2.
 DR TIGRFAMs; TIGR00657; asp_kinase; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
 DR Kinase; Complete proteome.
 SQ SEQUENCE 470 AA; 5046 MW; DD6FE74EFA2730C5 CRC64;
 Query March 24.7%; Score 695.5; DB 16; Length 470;
 Best Local Similarity 37.6%; Pred. No. 5,9e-45;
 Matches 164; Conservative 93; Mismatches 164; Indels 49; Gaps 16;

QY 85 VYMKRGSSVSSAAMAEVAGLILTFPEER---PVYVLSAMKTTNNLLAGEKAVGCG 140
 DB 2 VYMKRGSSVSSAAMAEVAGLILTFPEER---PVYVLSAMKTTNNLLAGEKAVGCG 57
 QY 141 VTH-----VSELEENMKSLIHKTVDELGLPKCNTSLY--ELEQLKGIAMKELTPR 193
 DB 58 CLEBAQVGEVRQCHLDLIGELISEBLOQEVIAKTEVLTRELRTEIEIVGELTER 117
 QY 194 TSDYLVSGECNSTRIFSAVYLNKIRYKARQVADPDIGFITTFDEFG---NADLIEATPA 249
 DB 118 SKDRFCSFEGELLSTSVFAALNEAGVSCMIDVETV-MTDDRGRFARPLAECQKNTSE 176
 QY 250 VAKRLHGMIDOPALPVVTGFLGKMGKSGAVTTTIGRGSDLTATTIGALGLREIQVMK 309
 DB 177 IIKPL-----LDAGTVVTVQYIG-ALEGRTTIGRGSDLSAALPFAMHSEIEIWT 231
 QY 310 VDGVLTCDPNIYPAKATVPVLTPEEATLAVFGAQLVHPQSMAPAREGDI PVRVKSYN 369
 DB 232 VDGVMTTDPRIYPEAKSIRVMTFSEAAELAVLGAQVLPDTIAPAVQKNI PVYVLTWHP 231
 QY 370 KAPGTLITPQRMDXGLV---VLTISYLVKSNVTMLDIVSTRMLGQVFLARVSGICIED 426
 DB 292 DSKGTLITDPPELLAGKSHGLVKSIAVKQAQALINIRSNMFGRHGFNSLPDV--FER 349
 QY 427 LCISVDCVATSEVSVSLDPSKTSWRELLIQASLHDHVEELEKTAIVYLLQQRAIISL 466
 DB 350 FGISVEMISTSEVSVSLTVDDAVV-SEPLIKALGALGV--EIE-----HKVATVSV 358
 QY 467 IG-NVEQSSLLIEKTRGLARK--SGVNVQMIISQASRVNMSLIVHSDAKALVHALQAFF 544
 DB 399 VGDNLKMSKGV--AGRIFNSLRNVNLMISQASEINVGWVDESVDQAIVSLHCEFF 455
 QY 545 EDDVLSQVEA 554
 DB 456 AE---SQGDA 462
 RESULT 13
 Q8T230 PRELIMINARY; PRT; 467 AA.
 AC Q8T230;
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Aspartokinase.
 GN LYSC OR MK0109.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 NCBI_TaxID=3320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AV19 / DSM 9639;
 RX MEDLINE=21927647; PubMed=11930014;
 RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.,
 RA Scharbardin O.V., Shakhova V.V., Belova G.I., Aravid L.,
 RA Metale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
 RA Malyn A.G., Koonin E.V., Kozayvkin S.A.; Methanopyrus kandleri AV19
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 and monophyly of archaeal methanogens";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL; AE010310; AAM01326.1;
 DR GO; GO:0016597; F:amino acid binding; IEA.
 DR GO; GO:0004072; F:aspartate kinase activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR InterPro; IPR005260; Asp_kin_monofn.
 DR Pfam; PF00696; aakinese; 1.

Search completed: March 23, 2004, 13:17:34
Job time : 77 secs

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Acroor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.B., Grahme D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.U., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanobrevibacterium smithii reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res 12:532-542(2002).
DR EMBL; AE010670; AM03585.1;
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:0004072; F:aspartate kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR001341; Aspartate_kinase.
DR InterPro; IPR005260; Asp_kin_monofn.
DR Pfam; PF00696; aak_kinase; 1.
DR Pfam; PF01842; ACT; 2.
DR TIGRFAMs; TIGR00657; asp_kinases; 1.
DR TIGRFAMs; TIGR00656; asp_kin_monofn; 1.
DR PROSITE; PS00324; ASPARTOKINASE; 1.
DR KINASE; Complete proteome.
SQ SEQUENCE 472 AA; 51022 MW; AFS130532FE1A5F7 CRC64;

Query Match 23.4%; Score 657.5; DB 17; Length 472;
Best Local Similarity 35.0%; Pred. No. 5,1e-42;
Matches 167; Conservative 95; Mismatches 182; Indels 33; Gaps 11;

QY 83 LSVYMKFGSSVSAPMAEVAAGLITPPE--RPVYVLSAMGKTTNNL-----LAGKA 136
DB 1 MKIVMKGGTSVDGKKIRVAQLLKYHEGNIIVVTGALGVTDRLLENARLASTKG 60
QY 137 VGGGVIVSEIEEM--NMVKSILIKTVDELGLPXICNTSLY-----ELEQLKGIAM 187
DB 61 -----KVLVKEFKTELTKHKEHAKDAIEDPRAKEVLYDLRIEELKALIGICVL 114
QY 188 KEITPTSDYLVFEGECMSTRIPSAVLNKIRKAROYAFDIFGFTTDFGNADILENTY 247
DB 115 GEITRSRIDYISYGERLAPIVSGAVRSIGASIEYTGEGAGIVTSDYGNARPLEKTY 174
QY 248 PAVAKRLHGMWIDDPAPVYTGFLGKMGSGAVTTIGRGSDLTATTIGKALGLREIOW 307
DB 175 ELVTKRL--GRLIS--HLVVTGFIGEN--EDGIITTLGRSGSDPSASILGALKADEIMWM 231
QY 308 KDYDGVLTCDPNTIYPAKTVPYLTFFEBATELAYFGAQLHPQSMREPAREGDIIVRYKNSY 367
DB 232 KEVNGIMTTPRIAPFEAKTIPOISYAEAMELSYFGANVLHPRIIEPAKREHIVPVKNYF 291
QY 368 NPKAPGTLITRQDMDEGLVVLTSIVKSNVTMLDIVSTRMLGOVGFLARVSGICIEDL 427
DB 292 NPEFGTLVVAEKQCHVAKAVSLI--KVALINISGABMPGTGTVALFTALARAQV 349
QY 428 CISYDVAATSEVSYSVLDPKIMSRELIOAGSELDHYVEELEKIAIVRLIOQRAIISLI 487
DB 350 NIWISQSSSESNLSFVVSSESHVSAALKALHAEFNREIYVEIITSDRNV-----CVAVAV 403
QY 488 G-NVEGSSLLIEKTRVLKSGVNVQWISOGASKNMSLLVHSDAKALVEALHQA 543
DB 404 GAGAGTTPGVAKRVEGALGNSMINTIIMISQSSQYINISFVREGDFAFAVAKTLDHF 460
```